DOCUMENT RESUME

ED 231 120 EC 151 995

AUTHOR Sloane, Howard N.; Endo, George T.
TITLE Using Self-Instructional Materials to Train Parents of Young Handicapped Children in Solving Behavior Problems. Final Report.
INSTITUTION Utah Univ., Salt Lake City. Dept. of Educational Psychology.
PUB DATE Nov 82
GRANT G0079101344
NOTE 114p.; Some charts may reproduce poorly.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Affective Measures; *Autoinstructional Aids; *Behavior Modification; *Behavior Problems; *Disabilities; Early Childhood Education; Elementary Education; *Instructional Materials; Observation; *Parent Education; Program Development; Program Effectiveness; Program Evaluation; Reliability; Research Design; Validity

ABSTRACT

The 3-year project developed self-instructional programs and evaluated parent use of these programs (approximately 185 families) to improve behavior problems of their handicapped children, aged 3 to 9. The project's format included five goals (e.g., determination of the degree to which parents can treat behavioral problems without professional contact) and development of seven "cookbook" programs to solve problems such as tantrums and interrupting. The first year was spent developing logistics and assessment instruments for the program and included visits by trained observers to the homes of 75 families during three 1-hour sessons per week; however, no interpretable data were obtained due to high family attrition and low frequency of child behaviors. The second year included lower attrition due to use of parent-recorded behavioral measures but these measures had problems with reliability and validity. Positive change was reported by nearly all the 20 participating families. The third year included significant modification of parent report measures resulting in reliable and valid data, and completion of the study by 24 families, thereby providing 72 evaluations across seven different programs, with such results as 59 evaluations showing positive change. Included are substudies on attrition, parent charts, and numerous graphs. (Author/MC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

FINAL REPORT

Using Self-Instructional Materials To Train Parents of Young Handicapped Children In Solving Behavior Problems

ED231120

Grant No. G0079101344

Howard N. Sloane
George T. Endo

Educational Psychology Department
University of Utah

November, 1982

E 15 1995

TABLE OF CONTENTS

Chapter	Title	Page
Abstract		
One	Introduction, Overview, Objectives General Procedures	
Two	Telephone Contacts Who Did Not Participate	
Three	Subjects Whose Data Were Collected by Trained Observers	
Four	Year Two Self-Report Subjects	
Five	Year Three Self-Report Subjects	
Six	Year Three Mail-in Subjects	
Seven	Auxiliary Study	

Appendix

Supporting entities

Existing presentations

Projected presentations

Projected publications

Abstract

This is the final report on a three year project to evaluate parent use of self-instructional programs to improve behavior problems in young, handicapped children age 3-9. Approximately 185 families participated in the project to some extent; many others who did not meet the selection criteria were turned away after an initial contact plus possible referral to a service agency if indicated.

The first year of the project was spent developing logistics and assessment instruments for each program to be used by trained observers who visited each home for several hours per week. Families were run using this observer format during all three years, but valuable and interpretable data were not obtained from this format due to high attrition coupled with the low frequency of behaviors recorded during a one hour observation period.

In the second year a format using parent-recorded measures of behavior was explored. This resulted in acceptably low attrition rates, but the measures developed had problems with reliability and validity, and the data collected lacked internal validity, although nearly all of the 20 families who completed the program in year two (60 programs in a multiple baseline design) reported positive changes.

In the third year the parent-report measures were modified significantly and methods of obtaining parent/observer indices of reliability and validity were developed and implemented. These data indicated that reliable and valid data could be obtained from this parent report format, although not all measures were as high as desired. Using a multiple-baseline design across three different programs (behavior problems) 24 families completed the study, providing 72

evaluations spread across seven different programs. If it is assumed that the maximum possible (but not necessarily desirable) gain from such a program is the decline of an undesirable behavior such as fighting, tantrums or interrupting to a frequency of zero per 24 hours, or an increase in compliance for a program such as following directions to 100% compliance, the median percent of possible improvement across all families and programs was 46% of possible. Using the same measure, 4 of the 72 evaluations showed no change (within 15%), 9 showed a negative change, and 59 showed a positive change. This actuarial summation was used for ease of communications in a purely descriptive manner and was not used for any statistical analysis.

Of the 72 evaluations cotained from the 24 families in year three 47 seemed to indicate internal validity as defined by reasonable stability and lack of trends during baseline, changes within the multiple baseline design which appeared associated with treatment onset, reasonable between condition changes compared with within condition variability, and reasonable magnitude of positive change (not actually internal validity). These were assessed by visual inspection of computer graphics. Twenty-five did not show clear changes associated with treatment.

In addition, a small sample of seven families completed the study using the same format except that no observer visits for reliability/validity were conducted. These seven merely mailed their data in periodically. Families were selected for this format due to excessive distance from the university, behavior problems which seemed very minor on initial screening, and willingness to participate. Results were similar to the prior 24 families reported except less dramatic. This is possibly due to the "ceiling effect" created as many

children had very little problem behavior in baseline.

In addition, the report includes some sub-studies on attrition and clinical use of programs, and demographic data on families, children, degree and type of handicap, etc.

CHAPTER ONE

General Overview

Overview of Report

The goal of this project was to evaluate parent use of self-instructional programs to improve behavior problems in young handicapped children, age 3-9. Program effectiveness and a number of other variables were examined. Final evaluation assessed seven programs for problems such as non-compliance, fighting, and five others. In earlier stages of this three year project three additional programs were examined, making a total of ten.

All major evaluations were done using a multiple baseline design across 3 programs. Families were selected where the target child had problems such that three of the seven (or ten) programs being assessed were appropriate for use with the child. Baseline data were obtained on all three programs, and after a time-lag of approximately two weeks the first program was given the parent to implement. Data collection on all three continued, and after another time-lag the parent was given the second program to implement. After another time-lag the third program was similarly given to the parent to implement.

Various methods of data collection with respect to child and parent behaviors were used at various stages of the project. At different times data were collected by trained observers, by parents or by both, with gradually evolving procedures and instruments. Measures other than child or parent behavior were also collected, including an extensive intake inventory, measures of parent fidelity in carrying out programs, measures of parent ability to answer questions about the programs, consumer satisfaction measures, and others.

Three peripheral studies were also conducted. One was an analysis of dropouts in the first year of the project. A second was an evaluation of the use of selected programs in a counselor-directed group

setting, rather than self-instructionally. The third was an analysis of parent ratings of disciplinary behaviors as a function of program use and outcomes.

Data from all three years of this project will be presented under a number of different headings. These are:

Chapter One: (current chapter) Overview.

Chapter Two: Telephone contacts who were not selected as subjects or who declined to participate, and procedures used with such subjects

Chapter Three: Subjects in groups using trained observers who completed the study in years 1-3, including dropout subjects in the trained observer groups who completed an intake questionnaire but nothing else, or who partially completed treatment. A sub-study comparing dropouts and non-dropouts from year one observer groups.

Chapter Four: Year two self-report subjects.

Chapter Five: Year three self-report subjects

Chapter Six: Year three mail-in data subjects

Chapter Seven: Auxiliary study on school psychologist referral.

Appendices: Publications, presentations, and support.

Prior reports as well as the original proposal and the continuation proposals have detailed many of the procedures, as well as including in appendices materials related to these procedures. In the interest of readability and conciseness such items will be described only briefly and omitted from appendices in this report.

Project Overview

The goals of this project were (a) to assess the degree to which

parents of young handicapped children can treat behavior problems using self instructional programs alone, and the differential effectiveness of each program, (b) to obtain information on the generality and persistence of any changes produced, (c) to determine if there are any factors which predict in advance success or failure of this approach, (d) to determine within each treatment program which aspects are actually critical for parents to carry out if they are to be successful, and finally (e) to determine family acceptance of this approach. Goals (a) and (b) were met.

This project's content differs from many parent training projects in three major ways. The first is that the "treatment" did not involve any professional or other staff contact with families, neither with parents or children. Of course, there was contact for purposes of data collection, but not for what is usually called treatment. The second major difference is that the programs used do not attempt to teach parents any general principles or skills related to parenting or child behavior. Each program used is, to a degree, a "cookbook" or recipe for solving a particular problem. The third difference is that data, in the current project, were obtained in natural home situations, and the behaviors recorded were those usually of concern to families. Thus, no "contrived" situations were used to "set up" some specific behavior, the naturally occurring frequencies of behavior in the home setting were recorded. In addition, behaviors corresponding to the concerns of parents, such as "tantrums" or "interrupting" or "not following directions" were the major child behaviors examined rather than categories mainly of interest to researchers. Similarly, parents learned skills focused on solving a problem such as excessive fighting, rather than on developing defined categories of parent behavior, such as

reinforcing, although the programs did require behaviors such as reinforcing from parents.

Recruiting and Selection of Subjects

Initial Recruiting and Screening. Subjects were initally recruited through a number of sources. Brief notices concerning the project, which included descriptions of the programs included and information for interested parents on how to contact project staff, were sent to mental health clinics, various special education projects, parent groups, and other public agencies or groups. Several of these groups requsted that a project staff member talk to their staff, which was done. These agencies referred potentially interested families to the project, indicating to the parents a phone contact.

In addition, recruiting was conducted through public media. The University Public Relations Office prepared written releases and articles which were sent to radio and TV stations, newspapers, magazines, periodicals, and other media sources. This resulted in a series of articles, appearances on TV and radio talk or news shows, and frequent mention by media persons in written, radio or TV formats.

Parents who contacted the project via phone through any of these sources were provided with brief information conerning the project and what would be expected of them, and had any questions answered. During this phone contact families were screened to insure that the child in question was at least three years old and not yet nine, and was potentially handicapped. Handicapped was defined exceptionally widely, and included physical, intellectual, emotional and behavioral handicaps. In addition, families who could not be reached by automobile within 30 minutes from the university were excluded from some groups. For those who passed the initial phone screening, an appointment was made for an

intake interview.

Intake-Interview, Informed-Consent, and Releases

At the intake interview, parents were first provided with a detailed informed consent statement. This statement described in detail all procedures to which families might potentially be exposed, stated that participation was voluntary, that families could withdraw at any time, and indicated how data would be kept confidential. Before a parent was allowed to sign this form, they were given a multiple choice test which covered the main points on the informed consent form. All questions answered incorrectly were gone over by the interviewer with the parent, and then the parent was allowed to sign the form if they so desired. Following this, the intake questionnaire was completed. If other information, such as from a professional or agency, was available, parents were asked to sign forms authorizing release of such information to the project. In year one, if the child attended a setting such as a school, daycare center, etc., from which generalization data might potentially be collected, those parents who so wished signed a release authorizing the project to contact these persons.

All families who completed these procedures were accepted as subjects.

Personnel and Training

There were several classes of project personnel, and the qualifications and training of each varied.

Observers. Undergradute and graduate students were used as observers when observer data were used. Several training procedures were assessed. The most recent procedure required (a) that the potential observer read and discussed written material on the observation procedure and codes, (b) that the potential observer then

observed and coded data from a standard videotape, (c) that the potential observer then conduct an "in vivo" hour long observation with a trained observer in a volunteer family who had completed the project and repeated this until satisfactory reliabilities were obtained. Most, but not all observers who collected the data reported in this study were trained in this manner. Others were trained more informally.

Transcribers and keypunchers. Undergraduate and graduate students were used as transcribers and keypunchers. Transcribers counted and recorded data from observer and other records onto a form suitable for keypunchers. Keypunch operators punched the data onto cards suitable for computer input. Both were informally trained by other staff.

Computer personnel. Computer programming and general layout of data handling by the computer was directed by a consultant from the University of Utah Computer Center. Some inputting of data and some programming was done by an undergraduate computer science student supervised by the consultant. Some programs were designed by an external consultant.

Managerial staff. Managerial staff consisted of two advanced graduate students (who varied from year to year) and the principal investigators, who were responsible for overall direction, design, and implementation. No formal training was provided to these persons.

Experimental Conditions

Baseline. During the baseline condition data were obtained on a program, but no treatment was in effect. Baseline data were always collected before the parent had obtained the written program. The length of time in baseline was determined in different ways for different groups, and was usually set at an arbitrary length of time. For initial groups using observer collected data, however the absence

of strong trends or extreme variability in the data, as determined by managerial staff who reviewed graphed data of the basic child behavior measures was an added requirement.

Treatment. After baseline was complete a family received the treatment program. Within a few days the family was given the "cognitive test" (described later) to complete and told to start treatment.

Post-treatment. At the completion of the third treatment the parent was asked to complete the consumer interview. At a later date in some groups follow-up measures were obtained after a varying length of time.

Additional Procedures in Some Groups.

After the intake interview, families in the trained observer groups indicated regular times which were convenient for observers to schedule appointments. Management staff assigned to each family an observer who had a schedule congruent with the family's. In self-report groups data collection procedures were described to families by visiting staff at this time.

As part of the intake, parents rated the severity of the behavior problems of the target child, and expressed preferences for program selections. This parent preference was the major factor considered in assigning programs. However, in observer groups some attempt was made to avoid assigning to families more than one program which relied heavily on parent collected rather than observer collected data, and to avoid assigning three programs which were difficult for observers to collect data on concurrently. In addition, combinations of programs which involved similar behaviors were avoided, such as whining and tantrums. If a parent did not like an assignment, it was changed.

Additional Procedures used only in Observer Groups

Observers contacted the parent and set up initial appointments. Observers were responsible for maintaining a schedule of appointments. Observers had to file regular "progress reports" which indicated the number of observations made and whether reliabilities were obtained. These were reviewed weekly by managerial staff. If less than five data collection sessions were held in two weeks, or if reliabilities were not obtained well in advance of when a condition might terminate, the observer was notified to schedule additional appointments or reliabilities.

During parts of years two and three observers were awarded cash bonuses for keeping to the appointment and reliability schedule. To obtain the bonuses the observer had to hand in progress reports on schedule. At a still later stage a bonus for families was added. When a family reached the point where it was likely that treatment on the third program would be initiated within two weeks, the observer brought them a letter signed by the principal investigators indicating that as a token of appreciation for their cooperation they would receive a $25 check when data were collected six times following the initiation of treatment on the third program.

Dependent Variables

Major outcome measures. The specific measures used and the procedures used to obtain data varied in different years. In the first year all measures of child and parent behavior were obtained by trained observers sent to the home. Starting in year two, families were added in which data were collected using ratings completed by parents. After preliminary assessment of these procedures in year two, in year three a revised parent rating method of data collection was used alone with

observer visits at approximately weekly intervals for reliability. In addition, in year three a small group of parents submitted data using parent ratings by mail, were never contacted face to face, and never had observer visits. In years two and three a small number of families in which trained observers were the primary source of data were continued, and in some of these occasional parent ratings were obtained for reliability. In each chapter the details of procedures used with the group being reported are included.

Cognitive Tests. Each program had a ten item multiple choice quiz which assessed familiarity with the procedures described in the written program given parents.

Consumer Satisfaction. A scale to evaluate satisfaction with the program and procedures was available in a parent and a child form. The parent form was completed in writing by the parent. The child form was administered orally to the child where possible.

Intake Questionnaire. The intake questionnaire was completed in the home by a staff member who interviewed the parent. The interview took between 1.5 and 3.0 hours. The questionnaire covered a wide range of demographic information, information on family constellation and structure, information on the child's handicap, and detailed information on possible target behaviors for both the project child and siblings in the 2-9 year old age range.

Independent Variables

Seven (ten initially) written programs, each of which described procedures a parent can use to correct a specific behavior problem with no professional assistance, were the independent variables. The seven booklets were taken from a commercially available publication (Sloane, 1979). The programs covered the following behavior problems or areas of

problems:

Following Directions

Whining

Fighting, Arguing and Teasing

Tantrums

Interrupting

Bedtime Problems

Room Neatness

Three programs used initially (mealtime problems, dawdling, chores) were dropped. The reasons for dropping these programs were a low rate of selection by families, behaviors which other programs could cover, and logistic problems in obtaining data.

Each program covered a relatively discrete behavior problem. Although each program varied in content and format, most started by providing material to help the parent decide if there was a problem, to define it objectively, and to determine if this program was appropriate for the child in question. The programs provided relatively step-by-step procedures for a parent to follow, with little or no explicit general information on parenting or behavioral principles, although some programs summarized in non-technical language the behavioral principles used. The written material included all charts, forms, visual aids, game materials or other materials used in the program. For this project separate copies of materials the parent might use were provided so parents did not have to tear pages from the written program.

Where appropriate, programs included some ssessment and data collection by parents, and some programs suggested initial or continuing behavior change procedures for the parent or other adults or family

members which might contribute to the target child's behavior improving. Several programs presented options in terms of procedures which could be used and guidelines for selecting which options were appropriate.

The average reading level of the programs was the 7.6 grade level, with a range of 6.8 - 8.9. Programs were bound in booklets measuring 21.50 cm x 13.75 cm., and in this format the texts (exclusive of charts, graphs, cartoons, etc.) ranged from 7-17 pages.

CHAPTER TWO

Telephone Contacts Who Did Not Participate in the Project

Many persons phoned the project to inquire about possible participation and to obtain more detail on what participation entailed. Some of those who inquired did not have a child of the correct age, a child who would qualify as handicapped, or a child for whom the programs being evaluated appeared appropriate. Such families were not accepted as subjects. After this phone contact other families decided for reasons of their own not to participate in the project. As most referral sources and media information for self-referral went into detail on project requirements and problems covered a large number of families who contacted the project seemed appropriate for an intake interview and desired to pursue participation after this contact. The number who terminated after this first contact was approximately 65, although this number may be inaccurate as many casual inquiries did not result in any written record for tracking.

All initial inquiries were handled by two of the managerial staff. Other staff who received a call took down a name and phone number, and the call was returned by the two managerial staff responsible for this task. This was to insure accurate and consistent contact information.

Staff responsible for this initial contact were familiar with community referral sources and experienced in dealing with families. Most families who appeared inappropriate or uninterested during this initial contact were given the names of several possible referrals (at their request), such as community mental health clinics or facilities available though the school districts. Many were also referred to the book from which project programs were taken - especially those who

appeared to have a delimited behavior problem and/or those who indicated they would not consider any sort of professional referral.

CHAPTER THREE

Families in Which Major Data Were Collected by Trained Observers

In the initial proposal the intention was to collect data from families by using trained observers who visited each home for three hourly sessions per week. Although this procedure seemed sound conceptually it did not lead to much useful data in practice. There were two reasons for this. The first was that parents found it onerous to maintain the schedule of observer appointments necessary under this procedure, approximately 40 appointments over a three month period. Of approximately 75 families who completed intakes and who were targeted for this group only about a dozen finished with sufficient data for analysis, and most but not all self-terminated for the reasons described.

A second problem also made the observer procedure less tenable than planned. Many of the problems cited are considered serious by families with frequencies of less than 10 per hour. Indeed, many had frequencies below five instances of the behavior per hour. Frequencies of one serious temper tantrum per hour or one serious episode of fighting per hour are quite disruptive, and amount to a significant amount of troublesome behavior per day. However, such frequencies generated problems for the research project. With baseline data for undesirable behavior between zero and one or one and two occurrences per hour positive effects due to treatment (reductions) are difficult to demonstrate. Yet social and economic factors made it impossible to observe for a longer period, and experimental and social restraints made it impossible to "set up" higher frequencies. In addition, with such low frequencies it was impossible to demonstrate reliability.

In spite of these problems, families were run using the observer data collection procedure and enough of them completed the program to initially warrant considering analysis. The additional procedures used with the observer families are recapped briefly below; details of procedures and copies of materials have been submitted in the first year final report and the continuation proposal for year two.

Measures in Observer Families

Major Outcome Measures. For most programs, major outcome measures were collected by trained observers using 30 second interval observing. Each of the programs had an observational scale designed specifically for it. All observational scales included a method of coding those intervals during which the behavior could not be observed (e.g., child ran in other room or bathroom) and inervals during which there was "no opportunity" to observe the behavior. Coding intervals as "no opportunity" varied with and was defined separately for each program. For example, fighting, which was defined as fighting with other children, was coded as "no opportunity" during intervals in which no other child was present. Intervals scored as "cannot observe" or "no opportunity" were deleted in all analyses.

The observational scale for each program included measures of both child and parent behavior, and varied for each program. The child behavior measures in each case related to various target behaviors associated with the program. The parent behaviors were ongoing parent behaviors which were specified in the program, for example, praising, criticizing, prompting, etc., and/or highly program-specific behaviors, such as whether a direction given by a parent specified a concrete behavior, or whether a parent left the room after an instance of child tantruming. The detailed observer codes are available in the appendix

to the year one report.

The interval observing sheet for each program was in the form of a grid with columns numbered in intervals and rows labeled for each behavior to be observed. For each family a master observation sheet with the interval grids for all programs being observed in that family was duplicated with 30 minutes of observation on a page. The back of each form contained space for identifying data including family code, observer number, date, program code, whether each program was in baseline or treatment, whether this was reliability or main observer data, etc., as well as information indicating whether this was the first, second or third sheet of observation data for that visit. To time intervals, observers carried a small solid state timer about the size of a package of cigarettes with a cord and earplug, which gave a brief beep every 30 seconds through the earplug. Each observer had a schedule for recalibrating and recharging the timer unit.

Reliabilities were obtained by a second trained independent observer who concurrently recorded the same data. During reliability checks (and on some regular observer visits) a tape recorder rather than the "beeper" was used to indicate intervals. The tape was pre-recorded with a "beep" and the number of each 30 second interval, corresponding to those on the observer sheet. Both observers used the same tape recorder with two 3 m. extensions and earplugs. This avoided problems due to one observer being one or two intervals ahead of the other. Reliabilities on all measures were collected at least once while all three programs were on baseline, once while only the first program was on treatment, once while the first two programs only were on treatment, and once while all three programs were on treatment, although in many cases there were more reliabilities.

For several programs, such as room neatness and bedtimes, scheduling non-obtrusive observations was impractical. In these programs the major data were concrete facts recorded by parents, such as actual time the child was in bed with the light out, or whether the bed was made when the parent checked. A few programs involved behaviors which can be serious problems even though they occur at a low rate. An example of this is tantrums; one serious tantrum per day can be a problem. However, the probabilites would be against this single tantrum occurring during the hour observer visit. For these programs parent data supplemented observer records. No reliability was obtained on parent records.

Other Measures of Parent Behavior. Additional measures of the degree to which parents implemented procedures indicated in programs were obtained from "parent-fidelity" checklists. Included were items such as whether or not parents had charts posted, whether charts were up to date, whether points or other reinforcers wrere dispensed, etc., and varied with each program. Observers completed these forms during regular observation visits according to a pre-determined schedule for each program. In some of the simpler programs this was done only once, approximately a week after intervention started. Initial checks also assessed procedures parents were required to do only at program instigation, such as discussing program requirements, necessity and reinforcers with the child. In programs where parents were required to make several significant changes in such procedues, or where major program implementaton involved updating such forms, checks were made more often.

These informal measures often depended upon the observer asking the parent what had or had not occurred, although for items which leave a

visible product, such as a chart, observers checked these items themselves.

Results of Studies Using Trained Observers

Although approximately 70 families were recruited and completed an intake in groups using or in which it was planned to use trained observers to collect data, only 11 families completed the study in this format. A high proportion of the initial 70 were early families who either dropped from the study or were dropped after initial R & D on procedures was completed. A high percentage of this R & D in initial stages related to development of observer measures and revision necessitated by logistics, behavior observed, or inadequate reliability.

Additional subjects were dropped as it was noted during year one that in many families reporting behavior problems the observed rate of problem behavior in a one hour observation was too low for purposes of research. Baseline rates for behaviors such as fighting, whining or tantrums between zero and three per hour were insensitive to positive change as little reduction was possible, and in addition created problems in recording, computing and demonstrating reliability. Satisfactory reliability for all scales was demonstrated if instances with such low rates are omitted from calculations. Of the 11 families completing the project using this format none had data with sufficient recorded behavior to be worth interpreting. This is a major reason that other formats, described in chapters 4, 5 and 6 were explored.

For the reasons stated time and resources were not wasted preparing these data in figures, graphs or tables, as there is nothing worth reporting.

Attrition Study

Like other projects which have dealt with clients in terms of behavioral parent training or traditional therapies, the present project also experienced a high rate of attrition, approximately 50%. As a result, a study was conducted by Mahan (1982) to determine whether factors associated with dropouts vs. non-dropouts could be identified. Her study examined a wide range of information taken from the project's client intake form (described elsewhere) and from actual baseline data. Information assessed included demographics, nature and severity of target child's problems, siblings, lag time between intake and baseline assessment, and length of baseline prior to treatment intervention.

Thirty-one families were included in the analysis. The families were categorized into four groupings, consisting of those who dropped out before baseline (Group I, n=14), those dropped out during baseline (Group II, n=8), those who dropped out during the first intervention phase (Group III, n=5), and those who dropped out during the third intervention phase and/or finished the third phase (Group IV, n=5).

Of the many analyses conducted six emerged as significant findings. (These must be interpreted with caution because of the small sample size and high probability of Type I error).

It was found that whether families had their target children currently in other settings (e.g. day care, nurseries, special schools) was significant($chi^2 = 7.8$, $df = 3$, $p < .05$). For groups I through IV, the percent of families with their children currently enrolled were 28.6%, 50%, and 80%. Cleanly there is no linear relationship and thus no interpretation is given.

The groups also differed significantly on whether their children had previously attended other settings ($chi^2 = 24.00$, $df = 12$, $p <$

.02). For groups I-IV the percents were 81%, 100%, 66.7%, and 0%. There is an apparent trend here in that all groups who dropped out of the project had a high percentage of their children in other settings. This was not the case for Group IV.

There was also a significant difference among the groups in the number of hours their children had attended settings outside of the home ($F = 3.9$, $df = 3,21$, $p < .05$). The means for groups I-IV were 1.8, 5.5, 5.5, and 3.6. This finding was not interpreted due to the lack of a linear relationship between dropouts and hours attended.

Parents' rating of sibling tantrum of problem was significant among the groups ($F = 3.0$, $df = 3,21$, and $p = < .02$). The means for groups I-IV were 2.01, 1.2, 3.4, and 1.8. Also parents' rating of siblings interruption problems was also significant among the groups ($F = 2.99$, $df = 3,21$, $p < .05$). The means for groups I-IV were 2.5, 1.6, 3.4, and 2.0. Again interpretation of both findings is difficult due to an absence of linear trends between point ratings and attention.

An ANOVA performed on the average length of baseline (days) for Groups I, II, III and IV produced a significant difference ($F = 13.42$, $df = 2,17$, $p = .05$). The means for groups II-IV are 23.13, 47.70 and 63.73. Surprisingly the families who stayed the longest with the project had the longest baseline. If there is any apparent trend in the findings, it lies with the fact that Group IV had the longest baseline and that it had no children who had attended outside settings. One may postulate that this group of families may have not had any negative experiences up to that time and thus were able to show more patience with the long delay of the project.

The only consistent findings which followed a discernible trend occurred for both whether the target child previously attended another

setting and for the baseline findings. It may be the case that in the former, dropouts continue to do so because of lack of immediate and perceptible progress and/or a lack of patience on the part of parents. Unfortunately we were not able to followup on this group of families to get more information.

That there was a significant positive linear trend associated with the groups and the amount of lag time between baseline and first treatment is enigmatic.

Families whose children have never previously attended outside settings were also those who experienced the longest baseline period (approximately two months). One might conclude that there is a patience factor present in these families because the present project was their first attempt at seeking outside assistance.

CHAPTER FOUR

Year Two Families in Which Self-Report Measures Were Piloted

Due to the high attrition and low frequency data obtained with many families in which trained observers obtained data during an hour per day, in year two we explored using self-report data obtained from parents. Our rationale was that if frequent appointments were uneccessary attrition might be reduced. In addition, if parents reported on data using a rating scale the problems of low frequencies might be avoided. As it turned out we were correct on both counts; however, the rating scales created new problems.

The year two families run under these procedures were started merely to pilot the procedures and to explore self-report measures.These studies were under the immediate supervision of a doctoral candidate working on the project as a coordinator. Due to the exploratory nature of these studies only brief baseline data were collected. In addition, it was originally hoped that the ratings by "significant others" would serve to determine the reliability of parent ratings. In retrospect this was unsatisfactory. However, a mass of interesting data developed which finally served as the doctoral dissertation of the graduate student (Mcloughlin, 1981).

Design and Non-Experimental Procedures and Measures Common to Entire Project

The experimental design and the non-experimental procedures were approximately the same as with observer-data families. Families were recruited and given informed consent, given an intake, given cognitive tests and given final consumer evaluations in the same manner as for observer families. Twenty-one families initiated the study and 19

completed it. The same multiple baseline design across problems was used. The same treatment programs (experimental variables) were used. However, there were some changes in procedures and the measures of behavior change were, of course, different. A few additional measures and procedures were also used.

Pre-Experimental Procedures and Measures

"Knowledge of Behavioral Principles as Applied to Children" (KBPAC) Test. The 50-item four element multiple choice KBPAC test (O'Dell, Tarler-Benlolo, & Flynn, 1979) invited parents' responses to descriptions of child behaviors. A matrix indicating behavioral principles derived from established parenting texts (Becker, 1971; Hall, 1971a, Patterson, 1971; Patterson & Gullion, 1968) relating to each KBPAC question was derived. Parent responses to groups of KBPAC questions were coded to knowledge of particular behavioral principles in the matrix.

Absolute Ratings of the Severity of the Child's Problem Behavior. Ratings of the severity of the child's problem behaviors were calculated from respondents' ratings of each problem from four alternatives: "no problem," "mild problem," "moderate problem," and "severe problem". This rating was assumed to be a statement of the severity of the child's problem by comparison to a stable standard.

Relative Ratings of Change in the Child's Problem Behaviors. Ratings from respondents of the degree of change over time in the child's problem behaviors were also solicited. The rating categories included a "no change" category plus 3 categories each for relative improvement or decrement ("slightly better/worse" "moderately better/worse", "very much better/worse"). The rating asked that the comparison be made with the child's behavior during the previous week.

These ratings were assumed to compare the child to whatever rating was given in the prior week, and thus relative ratings did not have a stable standard but one that shifted according to improvement (or decrement) recorded in the previous week.

Arrangements for Selecting and Contacting Significant Others and a Teacher. Each parent was asked to elect two significant others and one teacher who might make once weekly behavior ratings of the target child's behavior. Release of information forms giving this consent were signed by a parent and duplicates mailed to the relevant other raters. The parent was asked to elect two rather than one significant other to allow for the possibility of a refusal to cooperate by one. Each significant other was to be a person who met with the child at least once weekly. Parents were asked to alert the significant others. "Teachers" were identified as any nonrelative professional or paraprofessional having a regular instructional role with the child. This group included, in addition to certified public school teachers, Sunday school religion teachers, preschool teachers, and the staff of licensed and nonlicensed child care establishments.

Following election by the parent and the mail receipt of a release of information document by the significant other, initial telephone contact was made with the significant other. All contacts with significant others were by telephone. No recompense was offered to significant others.

The teacher who had been suggested by the parent was mailed a package of information describing the project. Contained in this package was a duplicate of the parent-signed release of information, a description of the nature of the project's intent, a note inviting participation and behavior rating forms relating to the indicated target

child. The package and all subsequent mail contacts with teachers contained stamped, project-addressed envelopes. The letter inviting teacher participation incorporated a tear-off slip wich asked for an indication of the teacher's willingness to participate. In addition, it asked the teacher to indicate whether or not he or she wished to receive a copy of the text from which the programs were taken at termination of the study. All contacts with teachers were by mail. No other compensation was offered to teachers.

Teachers who failed to respond to the initial mail invitation were sent a reminder letter; those who did not reply to this reminder leter were sent a letter asking them to indicate with a signature on a prepared form that they did not wish to be involved. Each teacher who, having made at least one child rating and who subsequently failed to respond within seven mailing days (from the date of mailing out the forms) to a subsequent invitation to rate a child was sent a duplicate rating form, a return envelope, and a reminder letter. "Mailing days" did not include Saturdays, Sundays, and official school vacations. A second duplicate reminder letter and set of rating materials were forwarded after seven more mailing days when no reply was forthcoming. Teachers who three times consecutively failed to respond to an invitation to rate were not contacted again. Teachers who three times returned rating forms within the requested one-week period were sent a brief letter of thanks.

Experimental (Baseline) Procedures

The three behavior problems deemed by the parent at intake interview to be most serious were addressed as target behaviors. Each parent at intake interview identified three of the available program behaviors as reflecting at least a "moderate problem" for their child.

Baseline KBPAC. At the completion of the intake interview the parent was given the KBPAC test.

Baseline Absolute Behavior Rating. During the intake interview the parent was asked for an absolute behavior rating on the severity of each of the child's three target behaviors. A second baseline absolute rating on each target behavior was collected by telephone from the parent one week later.

The significant other first named by the parent at intake inerview was contacted by telephone two days after the intake interview with the parent. The significant other was asked for an absolute behavior rating on the target behaviors at that time, and also seven days later.

The teacher nominated by the parent was contacted following the intake interview and invited to make an initial absolute behavior rating on the child's target behaviors. This absolute rating and one solicited one week later were collected by mail.

Baseline Relative Behavior Rating. At the intake interview, parents were asked for a relative rating of the child's three target behaviors by comparison with the previous week. This request was repeated by telephone one week later.

At the initial telephone contact with the significant other, a request was made for a relative behavior rating on all target behaviors. This request was repeated one week later.

In the initial letter to the nominated teacher, a request was enclosed for a relative behavior rating on all target behaviors. This mail request was repeated one week later.

Intervention

Parents were given between two and five days to read each program booklet prior to being asked to implement the strategies. Parents were

visited before they formally incorporated these strategies into their child management practice and given a test of parent's knowledge of the program content (cognitive test).

Checks on parents' fidelity to program demands were completed approximately four weeks after each parent began treatment with each program. These checks were completed prior to the parent putting the following program into treatment in the case of fidelity checks upon phase one and two treatments; and prior to the consumer evaluation interview for the final phase treatment. Fidelity checks were conducted by a project worker at the family home and consisted of asking parents for responses to pepared questions, and requesting sight of charts and illustrative materials.

Parents were contacted once weekly. Each rating contact which did not coincide with a visit by a project worker to administer a test of knowledge relating to a program's content, or a measure of the parent's fidelity to program demands, was conducted by telephone. Two personal visits to parents were conducted prior to the first treatment, and one each at the end of the first, second, and third treatments.

Intervention Absolute Behavior Ratings by Parents, Significant Others, and Teachers. Parents, significant others, and teachers were aked to rate each of the child's three target behaviors once in baseline (the period before parents started any program), and three times during the period of intervention with the "absolute" measure of the severity of each problem. The ratings allowed four alternativies: "no problem", "mild problem", "moderate problem", and "severe problem". The ratings were collected by telephone from parents and significant others, and by mail from teachers, at the end of the first month of implementation for each treatment program.

Intervention Relative Behavior Ratings by Parents, Significant Others, and Teachers. Parents and significant others were contacted by telephone on the same day each week to collect a relative behavior rating. This rating recorded the degree of behavior change through a comparison of the child's three target behavior at the time of calling with the same behaviors one week earlier. Approximately four days prior to calling the parent, the teacher was mailed a relative behavior rating form. The teacher was asked to complete the form when it arrived at the school and return it that day. This system attempted synchronized completion of the relative ratings by respondents.

A Christmas vacation occurred during the cycle of the intervention. The treatment in progress for each family at that time was automatically lengthened by one week. This procedure resulted in one extra rating per family for whatever phase was in progress at December 25, 1980.

Post-Experimental Measures and Procedures

KBPAC. The KBPAC test was completed by the same parent of the child at the post-experimental stage as at baseline. In each case this was the child's mother. The post-experimental KBPAC test was given to the parent at the end of the final visit to the parent by the project worker. The parent was asked to complete the test without outside assistance and return it to the project office in a provided stamped addressed envelope,

Consumer Evaluation by Parents. At the final visit to parents by the project worker, the parents were invited to complete the consumer evaluation used in all phases (sometims modified) of this project. Questions were read to the parent by the project worker who in turn noted the parent's responses.

Consumer Evaluation by Teachers. Following completion of the

intervention, teachers were mailed a copy of the consumer evaluation form modified for teachers. They were asked to return the completed copy in a provided stamped addressed envelope.

Post-Experimental Procedures. At the end of the intervention immediately following the consumer evaluation, parents were given the opportunity to select as personal copies any of the seven programs not previously used in the treatment of thier child. Significant others and teachers were mailed a letter of thanks. For those teachers who had indicated a wish to be sent a copy of The Good Kid Book (Sloane, 1979) at the end of the study, this was also enclosed.

Forms and Materials. Copies of all forms and materials were appended to the third year continuation proposal.

Results

Descriptive Data from Intakes and Other Sources. Parents undergoing the behavioral self-instructional training incorporating parent-ratings of children's behaviors were not different from the larger population (total project applicants) from which they were drawn. All parents in such programs volunteer for child management training, and thus are often assumed to be reasonably highly motivated. Nevertheless, attrition rates of between 40% and 60% have been experienced by many previous parent training research projects with volunteers as well as in our year one observer families. This high attrition was not found with these "self-report" families. The only instances of "dropping out" (one parent and two teachers) were for reasonable and understandable circumstances (illness, moving).

This investigation, in common with many previous parent training interventions, primarily investigated mother training rather than the training of both parents. In over two-thirds of program implementations

only the mother was involved in using program strategies.

Parents were predominantly of the middle social class, educated beyond high school, and in the case of fathers, employed. Families were larger than the national average and generally intact. Children of the parent-rating study had less experince of special education than first year families. The families for the year two parent-rating study applied about one year after the initial announcement of the project in the community. It may have been that the parents most concerned with their children's behaviors volunteered first and were in the year one observer families. Later applicants for parent training thus involved less seriously "handicapped" chilren. Despite this, the respective parents offered similar evaluations of the intensity and extensity of the children's behavior problems.

All parents were able to select three of the ten alternative programs which addressed a problem behavior of at least a "moderate" level for their child. More boys were volunteered as problems than girls. Parents articulated clear preference for particular management programs, as reflections of their children's behavior problems. Across all parents, the programs directed at initiating compliance with parents' instructions (following directions) and reducing fighting, arguing, and teasing were the most popular.

Boys outnumbered girls three to one in the initial 21 families. Age range of children was 30 to 121 months, with a mean age of 5.7 years. All families but one included a mother and a biological or adoptive father. Fathers had a mean age of 35.9 years and mothers 32.9 years.

All fathers were employed full-time (this would probably no longer be true in this area), and five mothers were employed. Slightly less

than half (48%) of the fathers had professional, administrative, technical or managerial jobs. Mothers averaged 14.2 years of school and fathers 15.8, with a modal value for both parents of 16. Annual family income ranged between $1200 and $50,000 with a mean combined gross annual income of $22,000.

The number of siblings of the target child ranged from 0 to 11 with a mean of 3.1. Fifteen of the 21 families had attended an educational or child care setting, but only 1 currently attended a special educuation setting.

On the average, parents stated that they had considered the child to be handicapped or have a problem for 33 months at the time of the intake. Approximately half of the children were labeled "handicapped, diagnosis unclear" by parents (48%) and professionals (52%). Two of the 21 parents considered their child to have a learning disability.

KBPAC and Cognitive Tests. Parents increased their knowledge of behavioral principles duing the period of the inervention (from 55% corrrect to 75% correct). The learned principles appeared to reflect and be generalizations from information contained in the intervention programs. Nevertheless, the level of parents' knowledge of program content (cognitive test scores) did not contribute more than 7% of the variance in the increase to parents' knowledge of behavioral principles across the entire duration of the study. Such a lack of relationship may, however, be confounded by the narrow spread in total correct responses on the cognitive tests of knowledge of program content. The level of responses on all cognitive tests was uniformly high, between 72% correct and 93% correct. It appears demonstrated by the KBPAC results that parents do learn to generalize information, at least at a cognitive level, when exposed to strategies directed to specific and

discrete behavior problems. Since the generalization is from the specific to general principles, it may be a assumed that parents can also use information directed at the remediation of one particular problem, for problems in an identifiably separate behavioral domain.

Parent Fidelity. While parents demonstrated an acceptable level of knowledge of program content prior to implementing treatment, they did not rigorously adhere to program directions. Parents' fidelity to program demands was at a generally low level. In particular, charts and illustrative materials intended for data collection on child and parent behaviors were widely ignored, as evidenced by the fact that 38% of the charts were not used by any parent, and the mean use was only 17%. Those program suggestions which required the greatest amount of parent investment to enact or implement, appeared to receive the lowest rate of incorporation into parents' repertoires. However, most parents were able to describe procedures wich they had devised to personalize programs to their child, indicating that parents were able and willing to supplement the behavioral strategies recommended by program pamphlets. Parents also widely reported using the program strategies with siblings and target children. In only a minority of cases were the programs read by both parents during the period of treatment.

Consumer Evaluation Data from Parents. Parents offered generally favorable and complimentary statements regarding their experience of the intervention. A majority of parents indicated that each program in particular and the intervention in general was responsible for improvement in the target child's behaviors. Reported treatment "effects" were generally positive. All programs were, on the average, considered by parents to have been appropriate, useful, easy to implement, clear, and pleasant for both parent and child. All parents

judged the programs worthy of future use and most determined them as either comparable or peferable to available parent training alternatives. Parent comments at the final consumer interview indicated that over the period of the inervention, many gained both understanding and control over their own and their children's behaviors.

Consumer Evaluation Data from Teachers. Teachers were not involved in implementing the behavior change strategies, only in providing assessment ratings on the child's behaviors. They were not alerted by the project personnel which behaviors were in treatment or when treatment began. Perhaps due to the minimal task asked of teachers, the great majority continued their involvement to project completion, including making a consumer evaluation. Teachers reported that rating behaviors was not burdensome, on the whole improved their relationship with the target child and parents, and was beneficial to their instruction of other children. Many teachers found real difficulty in providing an assessment of behaviors which were apparently problems situation-specific to the home. Most teachers asked for greater involvement in the mechanics of the behavior change, although many also expressed surprise that the child was enrolled in such a program.

Behavior Change Ratings. Parents, significant others', and teachers' ratings of both the gravity of the children's problems and behavior change over a period of time were approximately parallel. Whereas parents judged the gravity of the children's problems more severely than significant others, who in turn judged the problems more severely than teachers, all reported similar general trends of improvement to children receiving treatment. The improvements were evident across the entire duration of the study and from just before starting a specific implementation to a period after it was implemented.

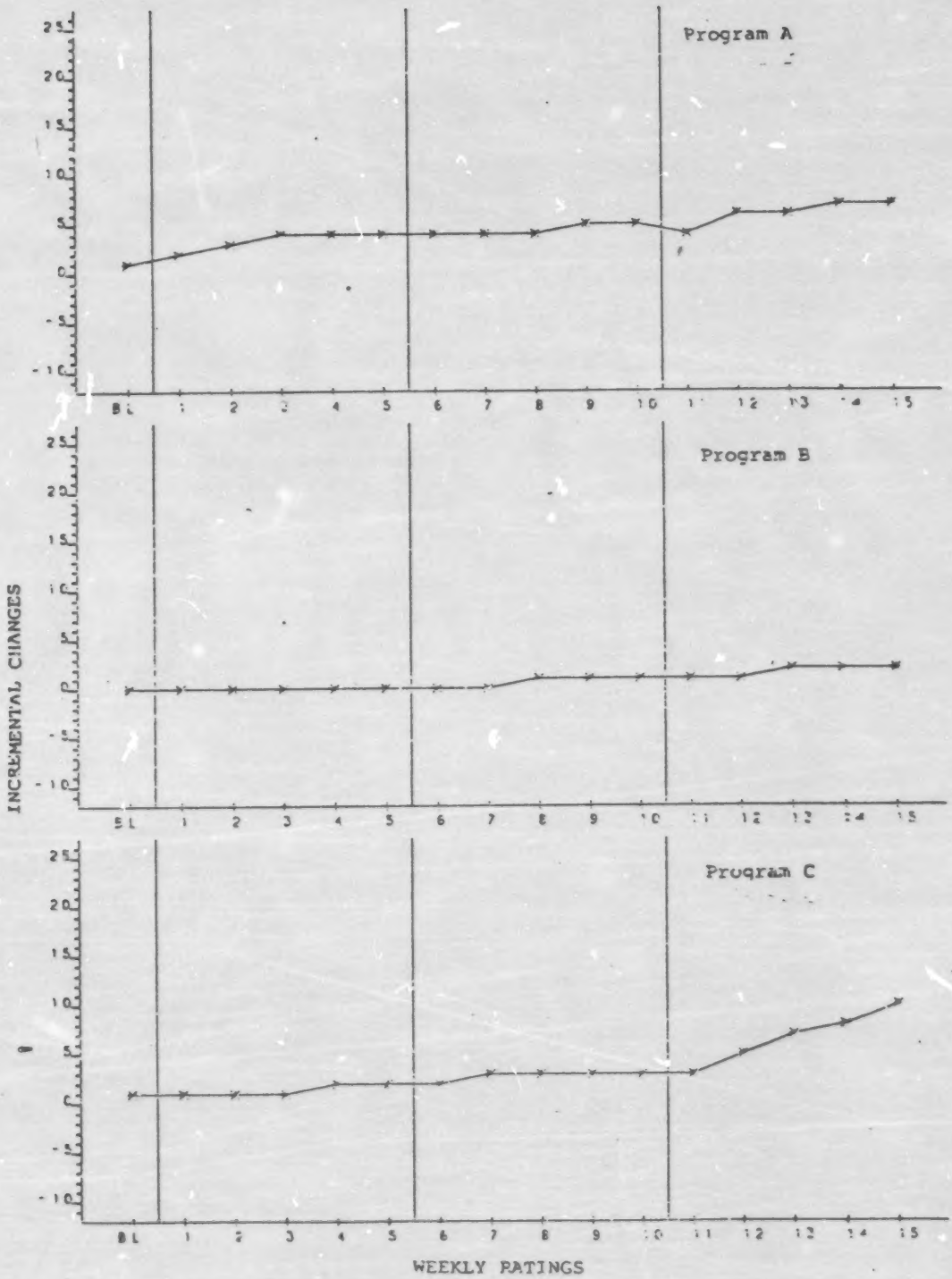


Figure 4.1 Multiple baseline for family code 45: dawdling (A); fighting, arguing, and teasing (B); mealtime (C).

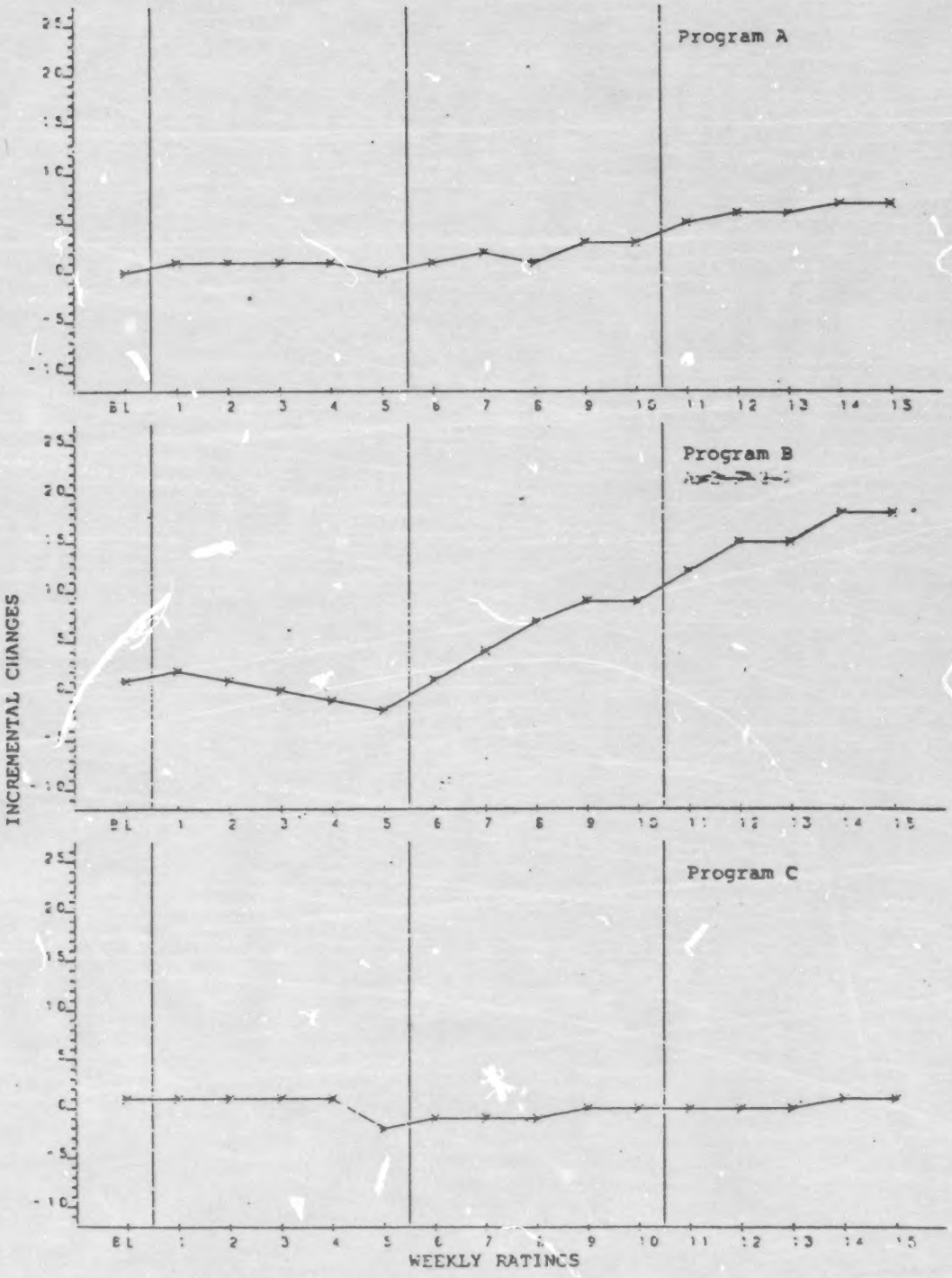


Figure 4.2 Multiple baseline for family code 4B: room neatness (A); bedtime (B); tantrums (C).

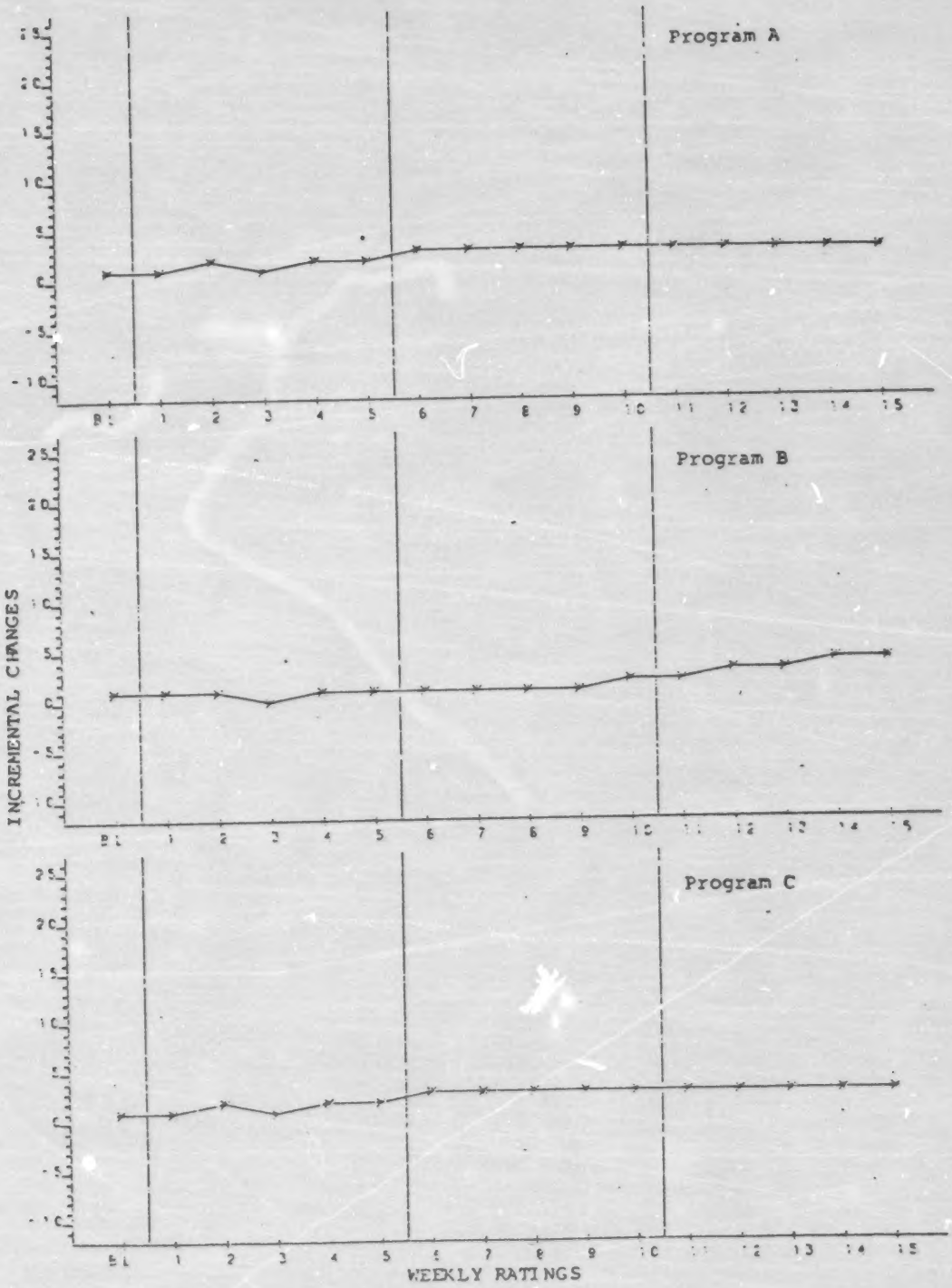


Figure 4.3 Multiple baseline for family code 50: following directions (A); fighting, arguing, and teasing (B); whining (C).

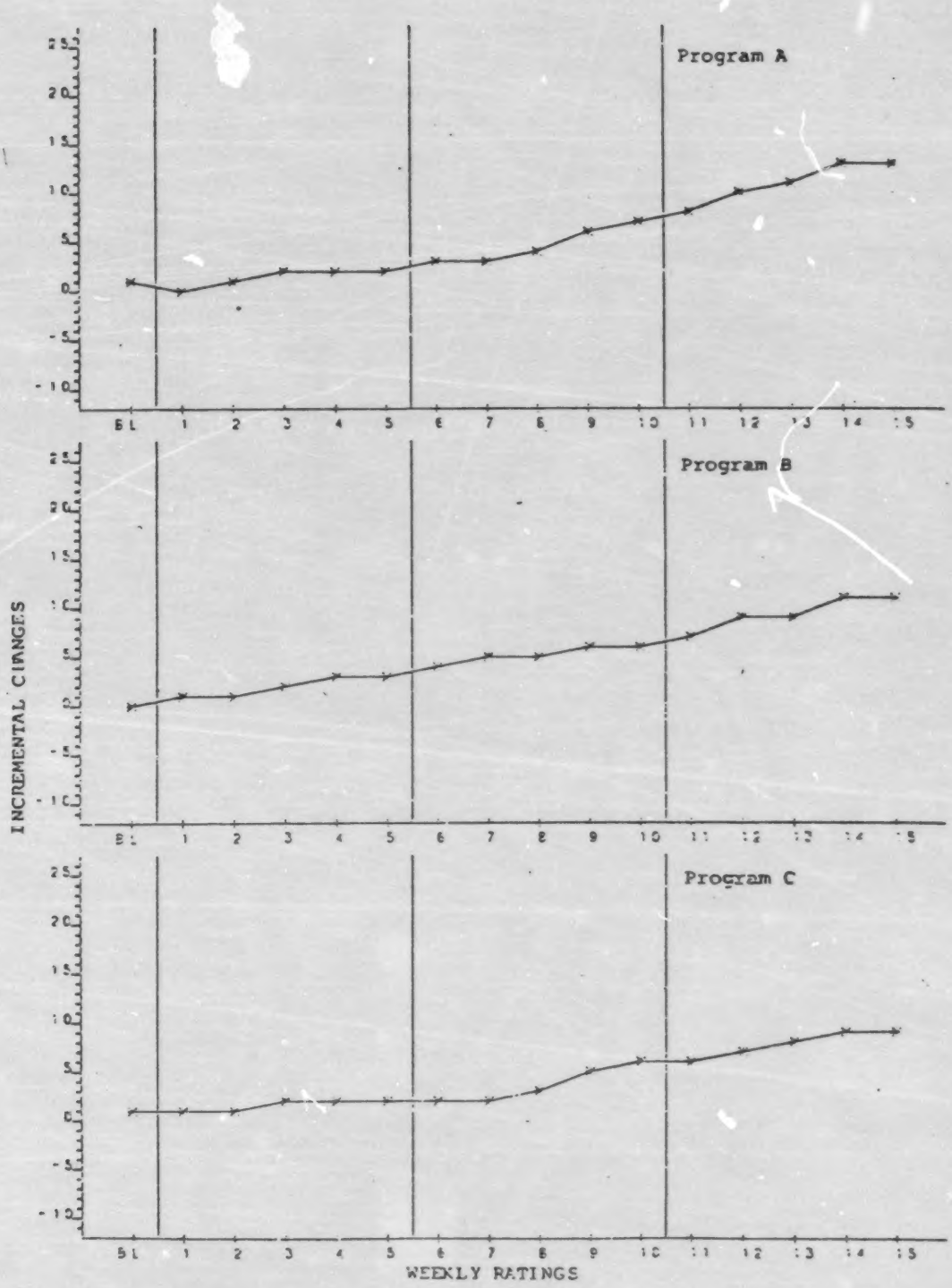


Figure 4.4 Multiple baseline for family code 55: following directions (A); fighting, arguing, and teasing (B); tantrums (C).

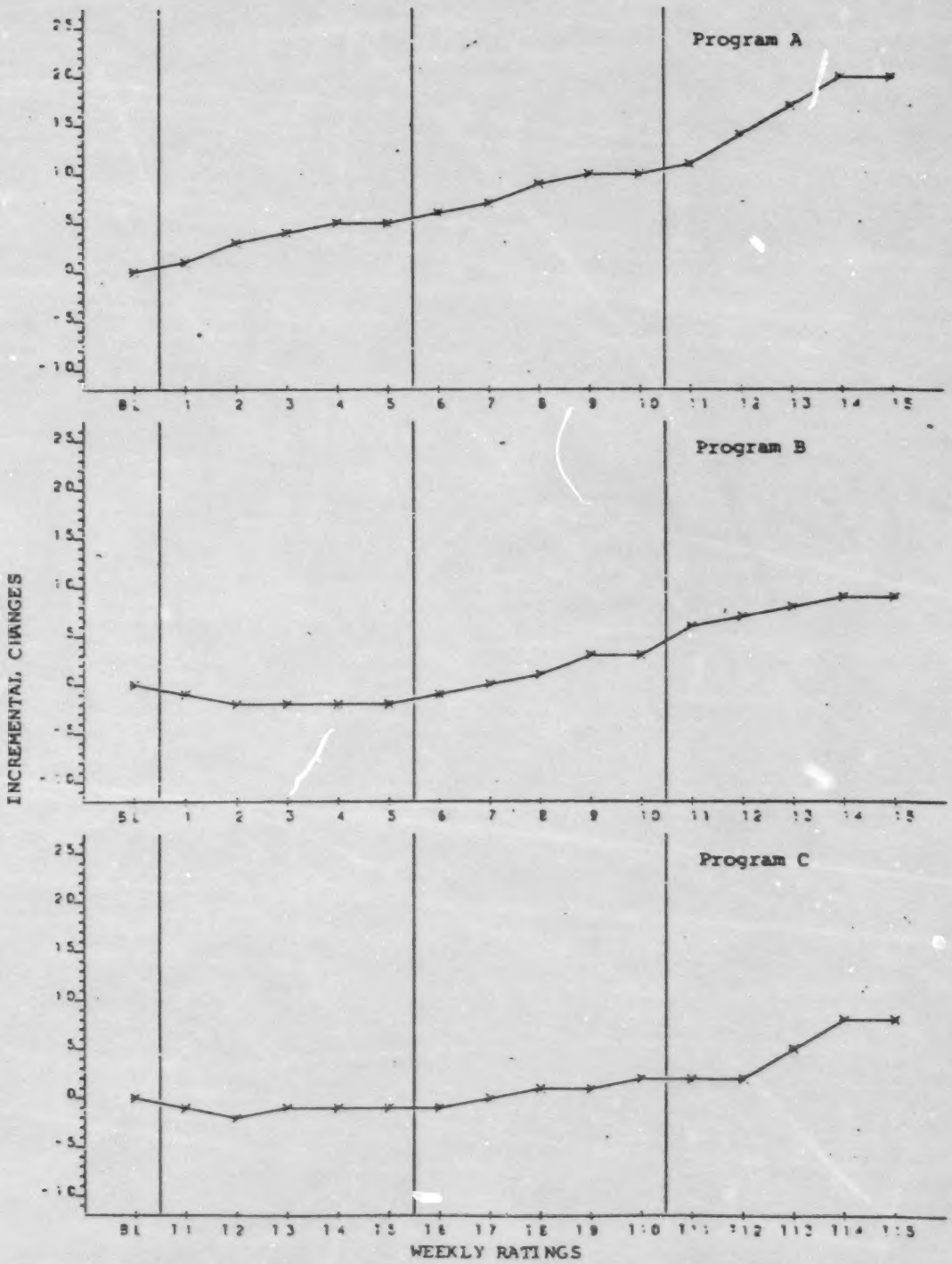


Figure4.5 Multiple baseline for family code 60: Mealtime (A); fighting, arguing and teasing (B); tantrums (C).

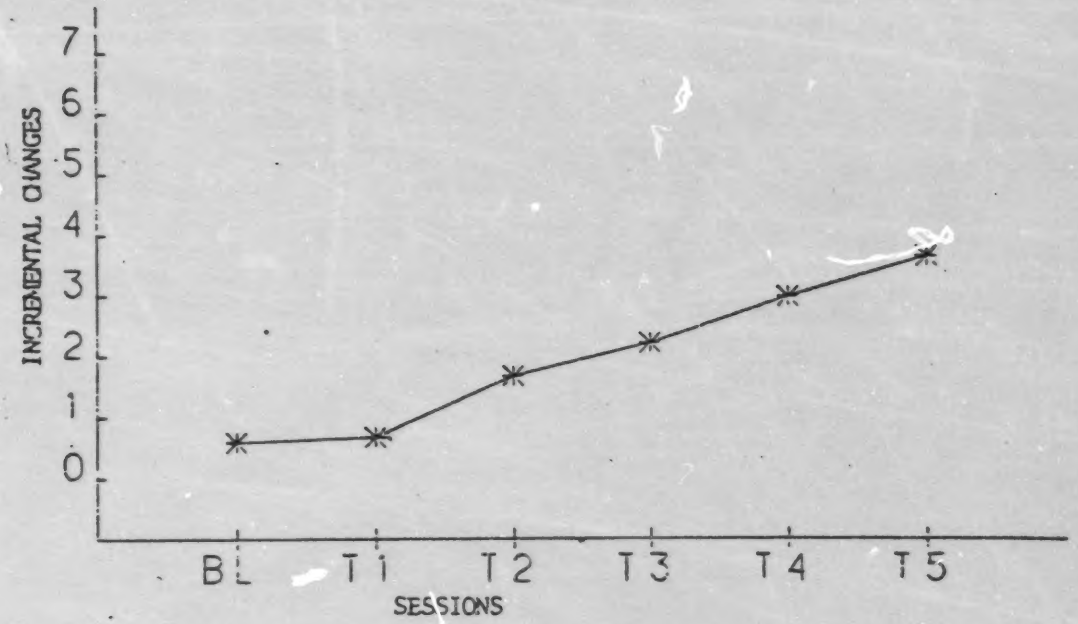


Figure 4.6 Mean baseline and mean aggregated treatment data across all phases for implementation of the following directions program.

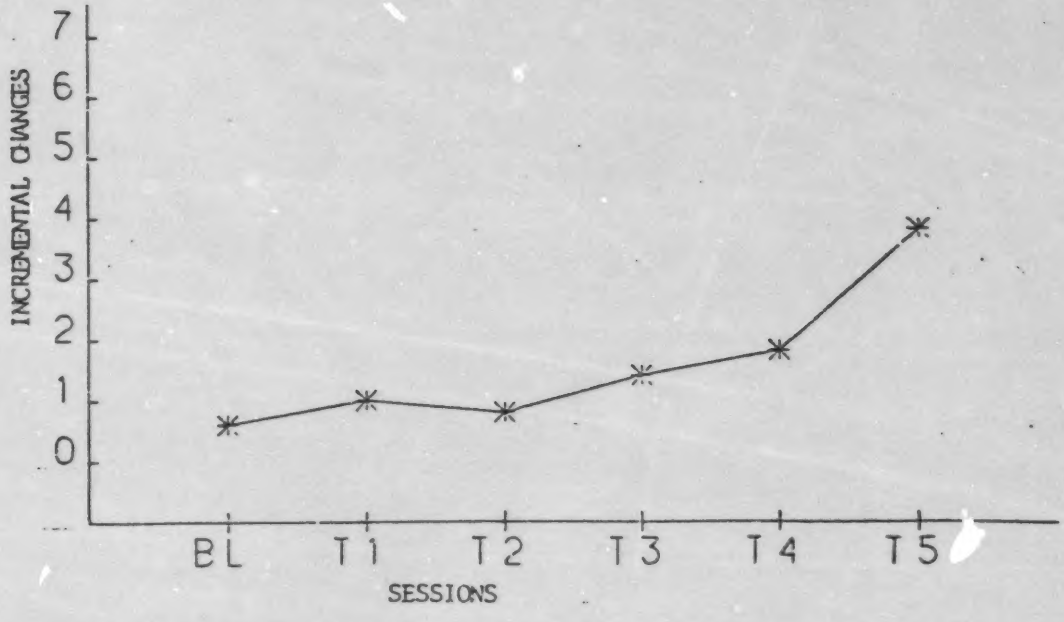


Figure 4.7 Mean baseline and mean aggregated treatment data across all phases for implementation of the whining program.

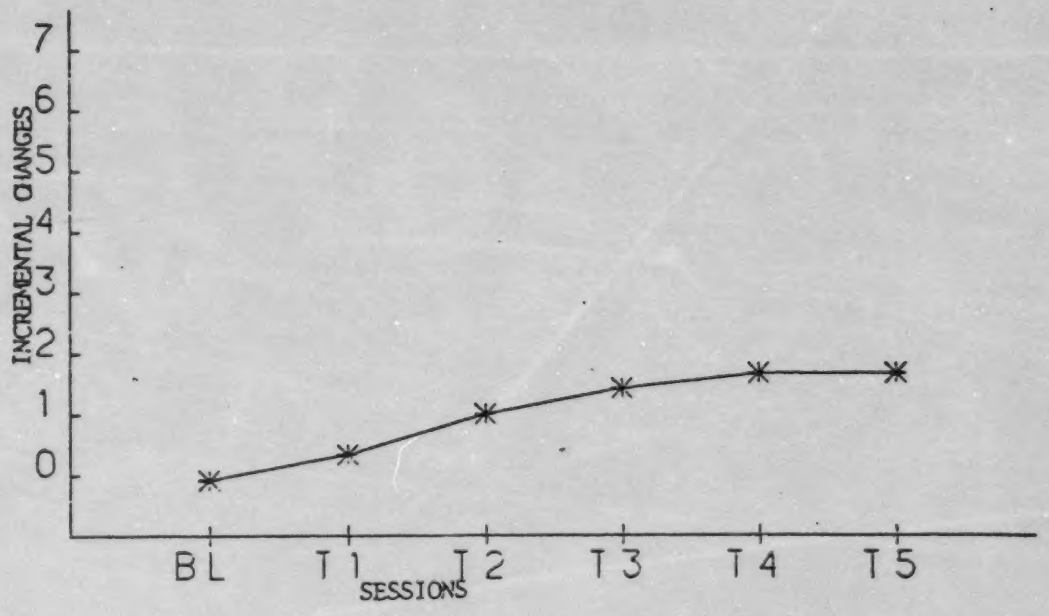


Figure 4.8 Mean baseline and mean aggregated treatment data across all phases for implementation of the fighting, arguing, and teasing program.

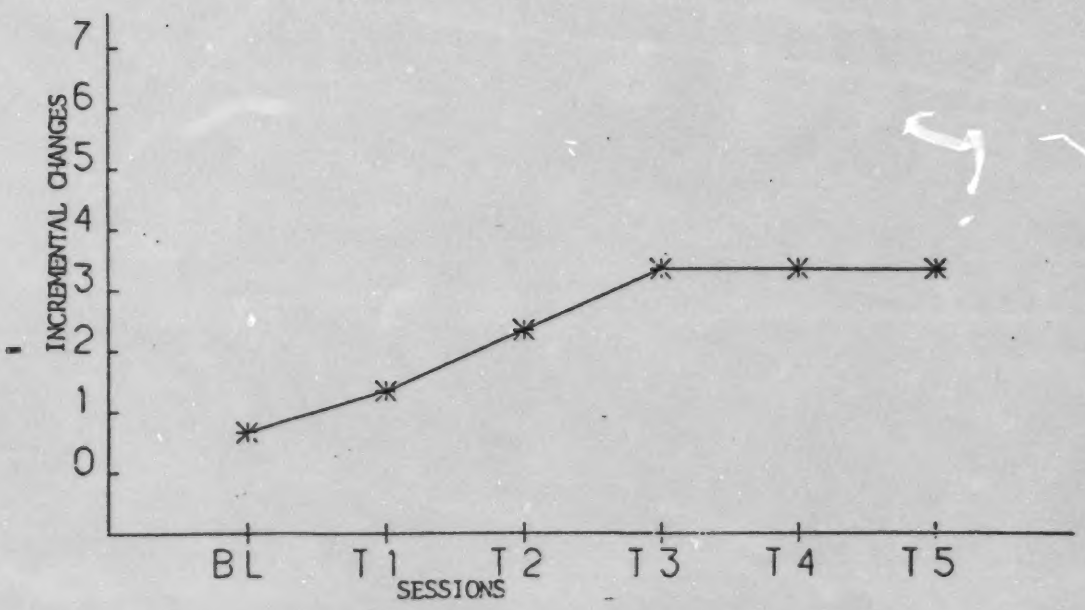


Figure 4.9 Mean baseline and mean aggregated treatment data across all phases for implementation of the dawdling program.

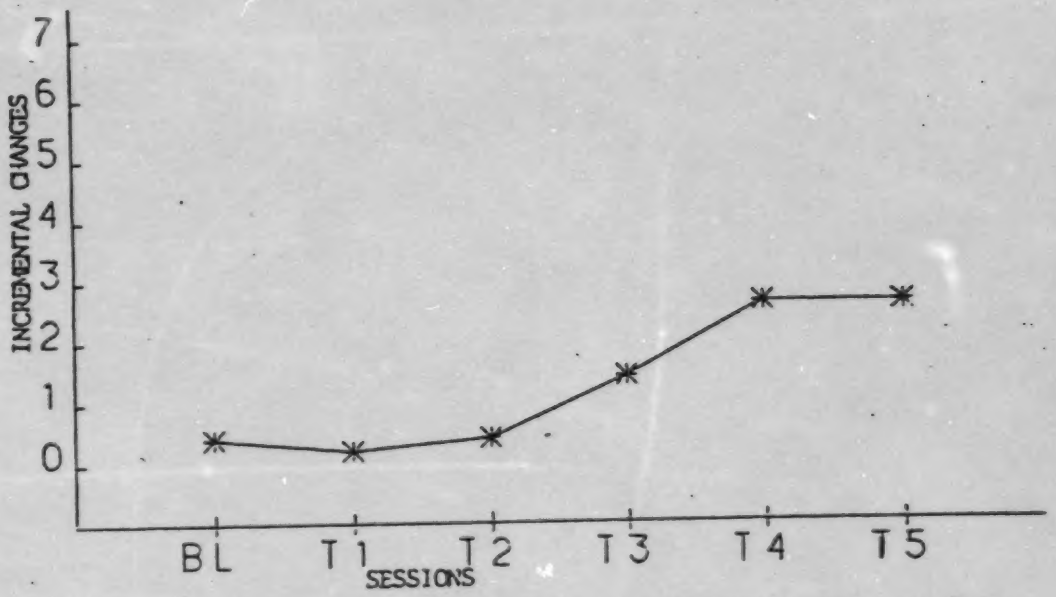


Figure 4.10 Mean baseline and mean aggregated treatment data across all phases for implementation of the tantrums program.

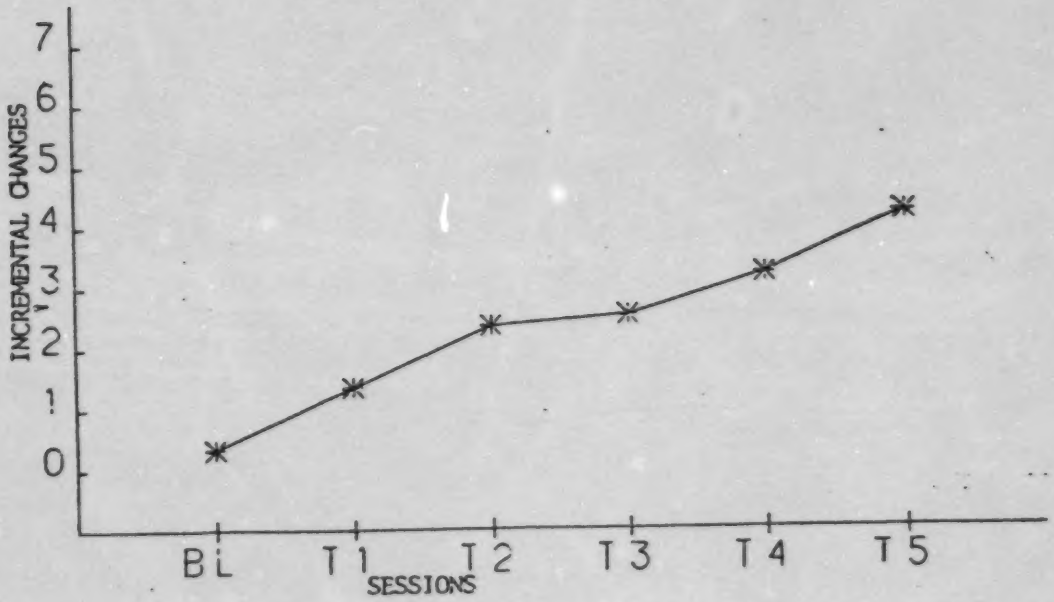


Figure 4.11 Mean baseline and mean aggregated treatment data across all phases for implementation of the mealtime program.

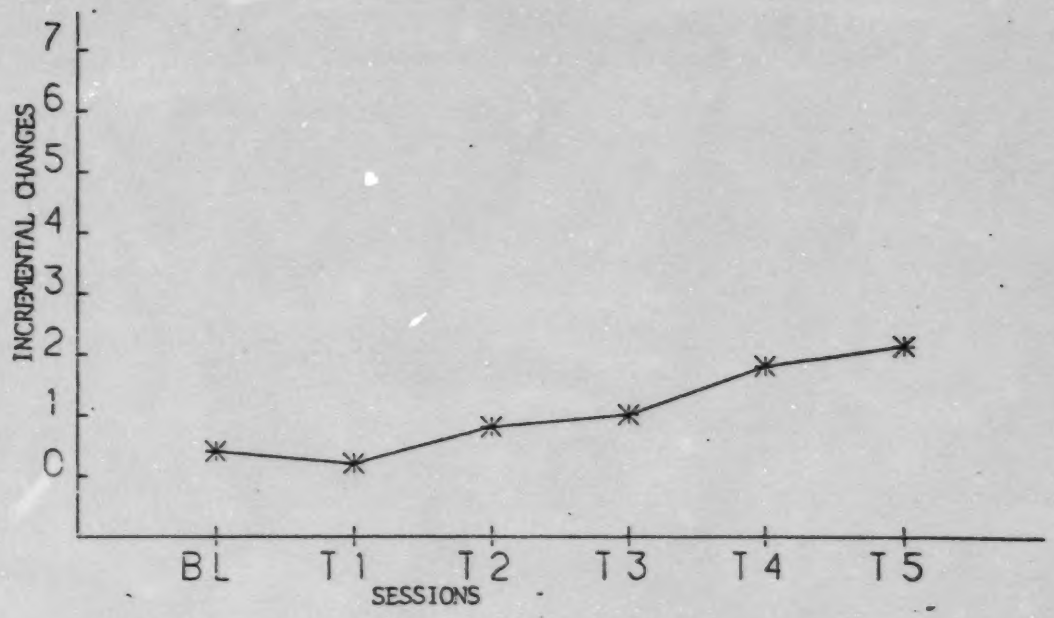


Figure 4.12 Mean baseline and mean aggregated treatment data across all phases for implementation of the interrupting program.

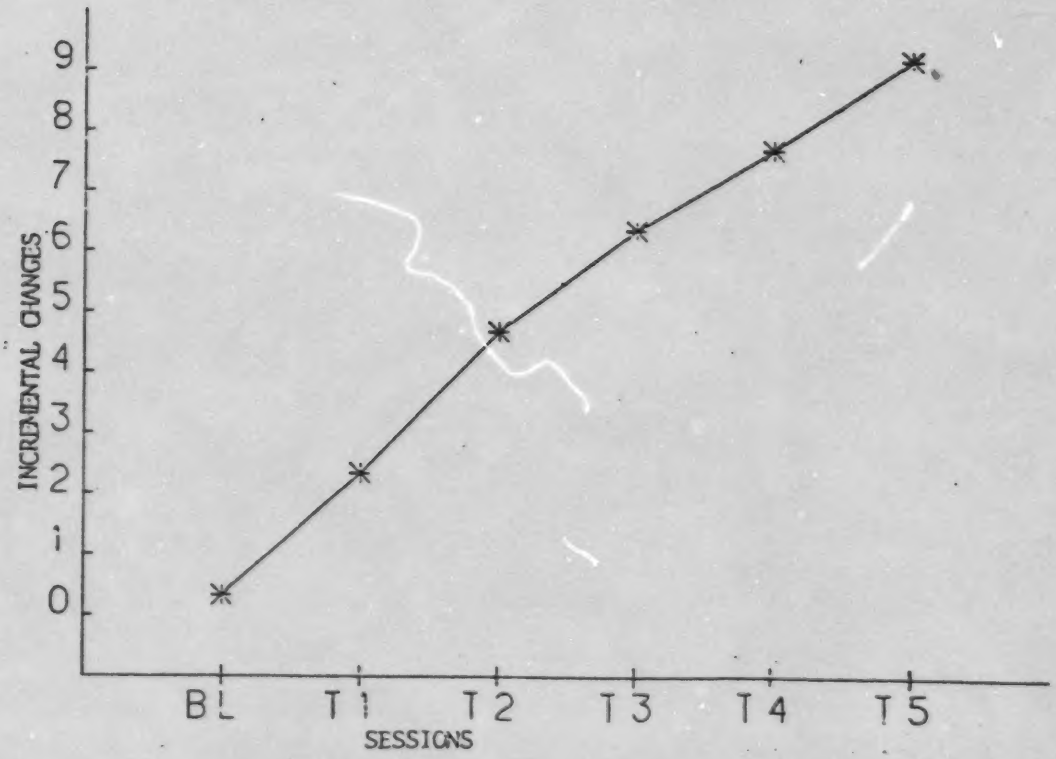


Figure 4.13 Mean baseline and mean aggregated treatment data across all phases for implementation of the bedtime program.

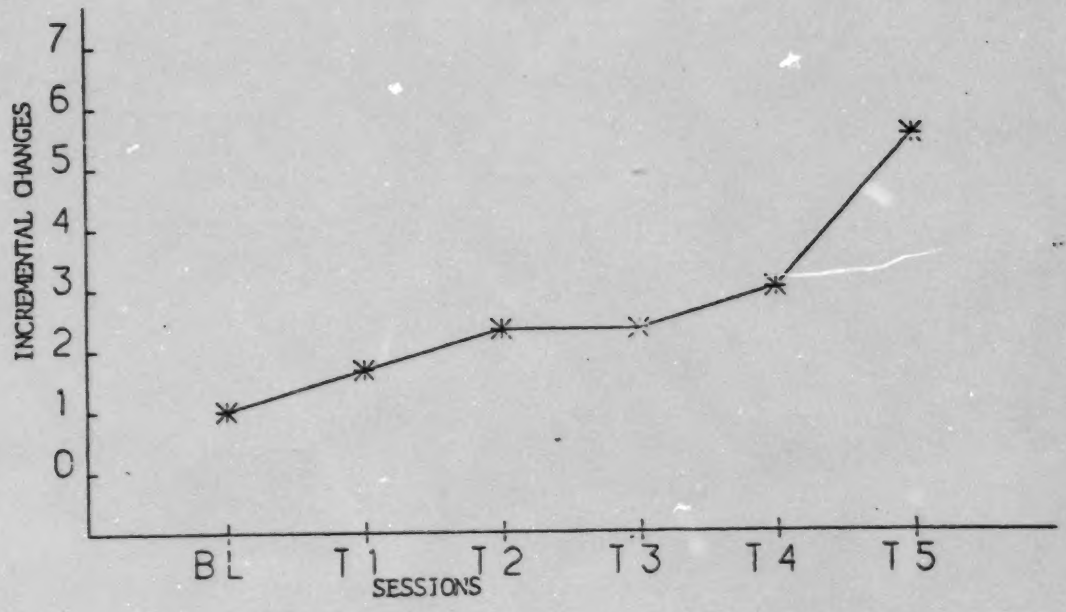


Figure 4.14 Mean baseline and mean aggregated treatment data across all phases for implementation of the grooming program.

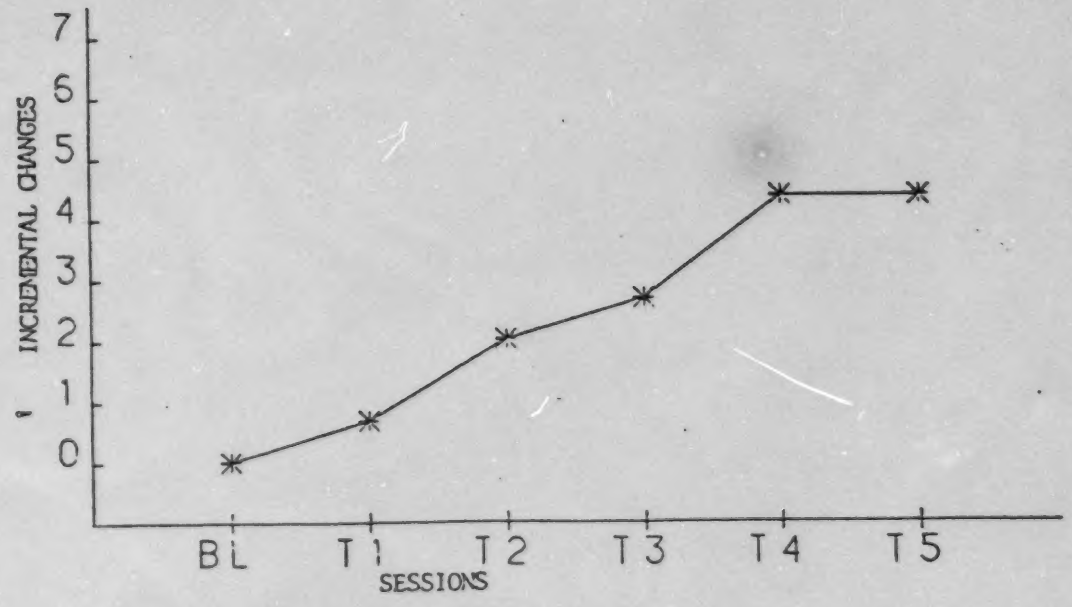


Figure 4.15 Mean baseline and mean aggregated treatment data across all phases for implementation of the room neatness program.

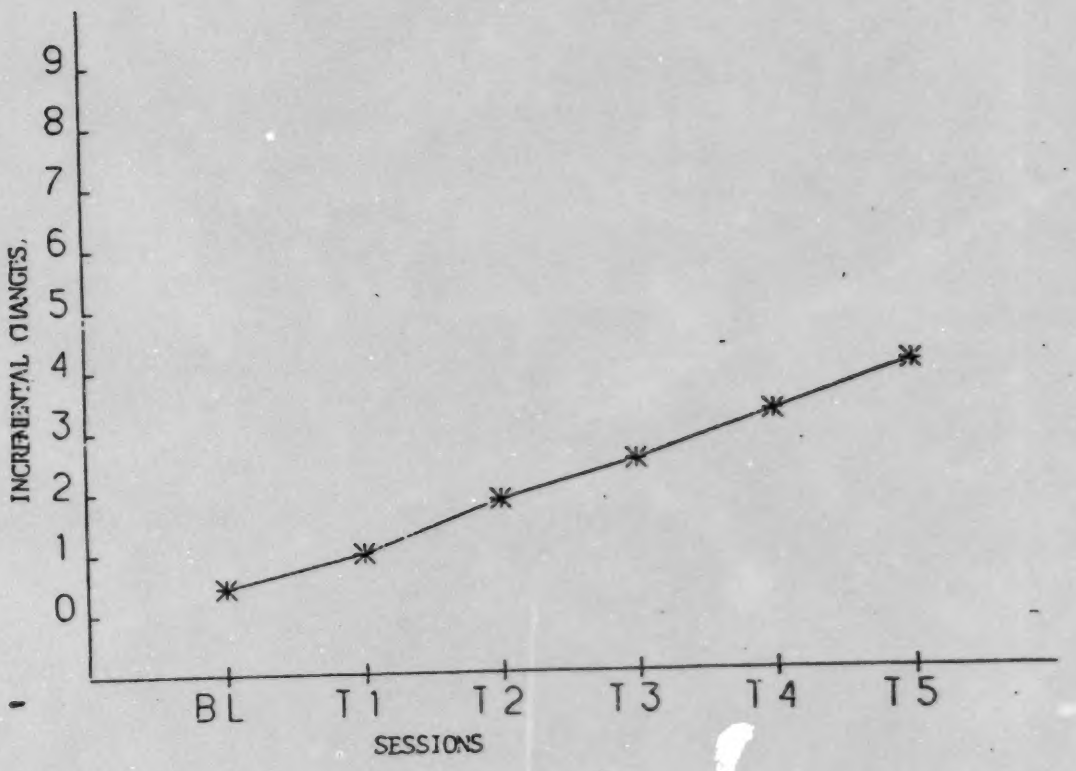


Figure 4.16

Mean baseline and mean aggregated treatment data across all phases for all intervention program data combined.

The existence or not of specific treatment effects for particular programs (internal validity) was undemonstrated.

Figures 4-1 through 4-5 present samples of individual behavior change for five families which are representative of the entire sample of 19 families. The data for the graphs are generated by giving a "no change" rating a weight of 0 (zero) and the other ratings a weight ranging from plus three to minus three. As the ratings are supposed to be made on the basis of change since the prior rating, each weekly rating is graphed from minus three below to plus three above the prior rating.

Figure 4-1 shows a family with very little change for two behaviors, and a change in the third (Program C) which started at the onset of treatment for that program. In 4-2, Program A shows gradual improvement starting some time after treatment for that program was implemented, Program B shows an abrupt and large improvement when treatment was implemented, and Program C shows a small decrement. In Figure 4-3 no significant changes are evident. In Figure 4-4 all programs show improvement, but in Program B it starts before treatment was implemented, as is also true for Program C. The same is true for Figure 4-5.

Figures 4-6 through 4-16 show the mean improvement for each program, regardless of whether it was implemented first, second or third. For this aggregation all treatments effects are grouped into 5 "sessions" which may represent one, two or three data points for a particular family depending upon whether treatment was implemented first, second or third. The aggregated treatment data show a positive effect for nearly every family.

These data, combined with parent impressions in the consumer

evaluation, indicate that parents clearly rated all programs as promoting improvement. However, the reliability of these data are unclear; their validity depends upon unknown reliability and upon whether parent opinion or objective behavior change are taken as the criterion or goal of parent training. However, even if satisfactory reliability and validity were pesent, the internal validity of the data is undemonstrated. The original plan to merely pilot a measurement resulted in only one baseline measure being taken. In addition, the fact that rated improvements did not have their onset coincident with treatment onset also prohibits causal interpretation.

Certain other "results" also need to be noted. Attrition was (approximately) only 1/5 to 1/10 what it had been in groups where data were collected by trained observers in the home. This obviously reduces costs. In addition, the actual costs of data collection by phone and mail were much less than training and paying observers to visit the home. All in all, it was estimated that research costs in this group were 1/5 to 1/10 the cost of data collection in observer groups.

In addition, although the range of scores was low in these families (using the year two self-report instrument) it was still generally higher than the 0-3 absolute range produced with observers in some programs.

For these reasons it was decided that in the final year we would explore self-report data collection further. We wished to do this in a setting with a better design (e.g. baseline), where reliability could be obtained, with measures that had more objectivity and range, and generally in a "tighter" way. This is reported in the next chapter.

CHAPTER FIVE

Year Three "Self-Report" Studies

The strategy for third year activities was based primarily on the use of the self-report measures used during the second year of the present project. The intent was to determine whether a self-report measure could be considered a valid measure of behavior change.

The third year activities initially involved sequential changes based on relatively immediate feedback pertaining to the degree of congruence between parent's self report and observer's recording of the behavior during any given hour. Failure to achieve acceptable levels of agreement in either the rating scales or methods of securing validation data or both.

The first change was made in the year two seven-point rating scale. Descriptions of each interval (e.g. mild, moderate) were eliminated and, instead, at intake, parents were asked to write an estimate of that frequency of behavior per hour which would correspond to the problem being severe to no-problem at all. The attempt was to have parent's form their own anchor points for the problematic behaviors. This rating was obtained initially duing intake and then once a week concomittant with an observer visit. Following each observer visit parents were asked to recall and write in an estimate of the number of behaviors they observed during the hour. This was then correlated with observers' recording. This procedure yielded promising but mixed results and was therefore ccontinued with additional procedures.

The final change consisted of having parents indicate with a nod or with a prespecified gesture that a given behavior had occurred. The observer then recorded parent indications or lack of it in addition to

INTERRUPTING

Instructions: Please fill in the daily rating boxes below. Your ratings should be an estimated average of your child's behavior based upon your interaction with the child on each day. Please fill in the ratings at the end of the day in order to reflect an entire day's behavior. The telephone data collecter will ask you for these daily ratings when they call.

A. What is your estimate of the average number of interruptions that occured this day?

B. What is your estimate of the average length of time in minutes of each of these interruptions?

C. What is your estimate of the average intensity of each of these interruptions? (use scale below)

1. hardly noticeable
2. noticeable, but no big deal
3. intense enough to be annoying, but not a big problem
4. on the border between no big problem and intense enough to be a big problem
5. intense enough to definitely be a big problem
6. very intense and desruptive--definitely a big problem
7. extremely intense--about as intense as possible--a big problem and quite disruptive

Family Code ___

Parent making daily ratings _____
(Mother or Father)

Child's Name _____

Week Number _____

Date ____ ____ ____ ____ ____ ____ ____

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
A.							
B.							
C.							

WEEKLY TWO-PARENT RATING

Instructions: Fill in the date and time of day start and stop in which both parents were with the child for the entire time period. Then make your ratings below in the same way as the daily ratings are made by one parent, but make these ratings independently of each other. Do this once each week on this form.

DATE _____ TIME OF DAY START _____ (am,pm) TIME OF DAY STOP ____

Rating	Daily Rater	Weekly Rater
A. Number	_____	_____
B. Length	_____	_____
C. Intensity	_____	_____

FOLLOWING DIRECTIONS

Instructions: Please fill in the daily rating boxes below. Your ratings should be an estimated average of your child's behavior based upon your interaction with the child on each day. Please fill in the ratings at the end of the day in order to reflect an entire day's behavior. The telephone data collecter will ask you for these daily ratings when they call.

A. (1) What is the average number of instructions given by you to your child each day?
(2) Of the instructions you have given each day, what is the average number which have been followed?

B. What is your estimate of the average speed with which your child followed those instructions he/she did follow?

1. instantly
2. promptly if not instantly
3. with a small amount of dawdling
4. with a moderate amount of dawdling
5. quite slowly
6. very, very slowly
7. it took forever--about as slow as humanly possible

C. What is your estimate of the average difficulty you had in getting your child to follow a direction?

1. no problem whatever
2. very slight difficulty
3. difficult, but not really a problem
4. a slight problem with moderate difficulty
5. quite difficult
6. a great amount of work--very difficult
7. about as difficult as anything I can think of

Family Code _ _ _

Parent making daily ratings _____
(Mother or Father)

Child's Name _____

Week Number _____

Date ____ ____ ____ ____ ____ ____ ____

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
A. (1)							
(2)							
B.							
C.							

WEEKLY TWO-PARENT RATING

Instructions: Fill in the date and time of day start and stop in which both parents were with the child for the entire time period. Then make your ratings below in the same way as the daily ratings are made by one parent, but make these ratings independently of each other. Do this once each week on this form.

DATE _____ TIME OF DAY START _____ (am,pm) TIME OF DAY STOP _____

Rating	Daily Rater	Weekly Rater
A. (1) Number	_____	_____
(2) Number	_____	_____
B. Speed	_____	_____
C. Difficulty	_____	_____

recording behaviors as usual. Additionally, parents were asked to recall (as previously) an estimate of the number of behaviors which they felt occurred during the hour. Thus validation data were obtained on an interval by interval basis as well as on an hour by hour recall basis.

Demographic and Intake Information

In the third year, 24 families served as subjects using the procedures described. Ten additional families were "drop-outs". For some analyses families which received the "recall" reliability procedure only are combined with those who received the interval or "call out" reliability procedure; for discussion of reliability and validity there is some separate discussion of these two sub-groups.

Of the total, 22 were two-parent families and the remaining two were single parent (mother-only) families. The mean age of the fathers was 33.5, the mothers 31.4. The average years of schooling for the fathers was 15.6 years, for the mothers, 14.8 years. All fathers were employed, 15 in professional occupations (educators, M.D.s, etc.) 2 in sales, 2 in crafts such as carpentry, masonary. No data were available for the remaining fathers. Of the mothers, 18 were housewives, 4 professionals, and 2 in sales. The mean income was $25,753, with a range of $12,000 to $50,000. The average number of siblings was 2.37 per family.

The average age of the target children in months was 67.8, with a range of 40 to 145 months. Of these 14 were males and 10 females. The majority had no apparent physical handicap and were classified as behavioral problems, three had specific handicaps as verified by physicians, one had multiple handicaps and two had undiagnosed handicapping conditions.

Most families (20 of 24) were sending their children to outside

settings such as nursery school or preschools at the time of participation. Sixteen families had sent their chilren to outside settings prior to the study. However, most were not sent to special education programs.

Cognitive Test Scores

The mean of the three cognitive test scores was 6.5. The maximum score possible was 10.

Consumer Evaluation Results

Nineteen of 24 families completed the consumer evaluation. Several questions were asked to assess the merits, if any, of the programs. When asked to give an overall rating of the programs used, 15 families rated the programs as "good" and 4 rated them as "neutral". (Each program was rated more than once since several families used identical programs.)

When asked if the programs produced positive changes in them (parents), all 19 responded "yes"; when asked if positive changes were produced in their child 17 of 19 respondents answered "yes"; when asked if positive changes were produced "in general" 18 of 19 parents responded "yes"; and when asked if programs produced positive changes in others (e.g. siblings, relatives, peers) 17 of the 19 responded "yes".

When asked if the programs produced negative changes in them (the parents) all responded "no"; when asked if negative changes were produced in their child 16 of 19 reported "no"; when asked if negative changes were produced in general all 19 parents reported "no"; and when asked if negative changes were produced in others 15 of 18 reported "no".

When asked if they would want to use programs such as these again, all 19 parents responded "yes".

A five point rating scale was developed to assess parent's perception of program usefulness, ease of use, clarity of programs, and whether the programs were unpleasant to use for parents and for chilren. For the usefulness and ease of use dimensions, low scores signified positive parent perception, while high scores signified negative parent perception. The reverse was true for the clarity and unpleasantness dimensions. Parents rated each of the three programs used.

For program usefulness, Programs 1 through 3 received ratings of 1.37, 1.73, and 1.89, with an overall rating of 1.65. For ease of use, the ratings for each were 2.37, 2.37, and 2.74, with an overall rating of 2.49. For clarity, the ratings for each were 4.52, 4.58, and 4, with an overall rating of 4.35. For unpleasantness of use for parents, the ratings were 3,84, 3.74, and 3.5, with an overall rating of 3.70. In terms of unpleasantness to children, the ratings were 3.63, 3.47, and 3.47 with an overall rating of 3.52. In general, parents perceived the programs as positive or somewhat positive on all dimensions.

Year Three Outcome Data

As we have no reason to believe that variations in the reliability and recording procedures would affect how well parents carried out programs, all year three outcomes are reported in this section, except the "mail-in" families for whom no reliabilities are available. These are reported separately in a brief chapter following this current chapter. Twenty-four single-subject replications were completed and reported in this chapter, leading to 72 program evaluations.

Each of the 24 families provided a complete multiple-baseline design across three programs. Figures 5-1 through 5-24 present the data for these 24 replications. The table which precedes the figures provides the key to the programs presented and the measures used.

KEY TO GRAPHS

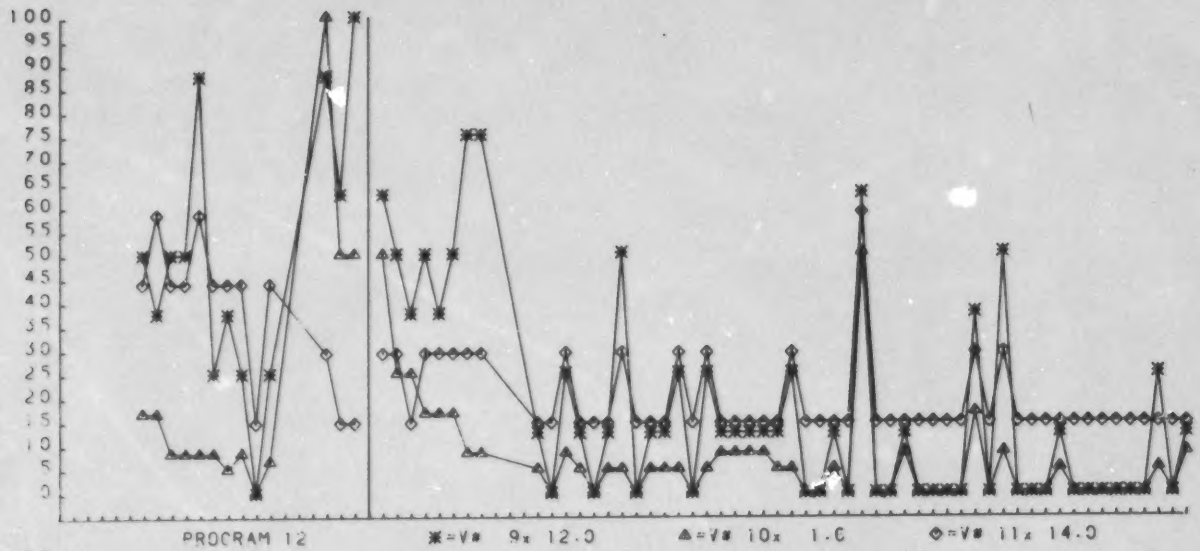
Family Identification: The Family Number used for confidential identification is given in the lower left hand part of each figure.

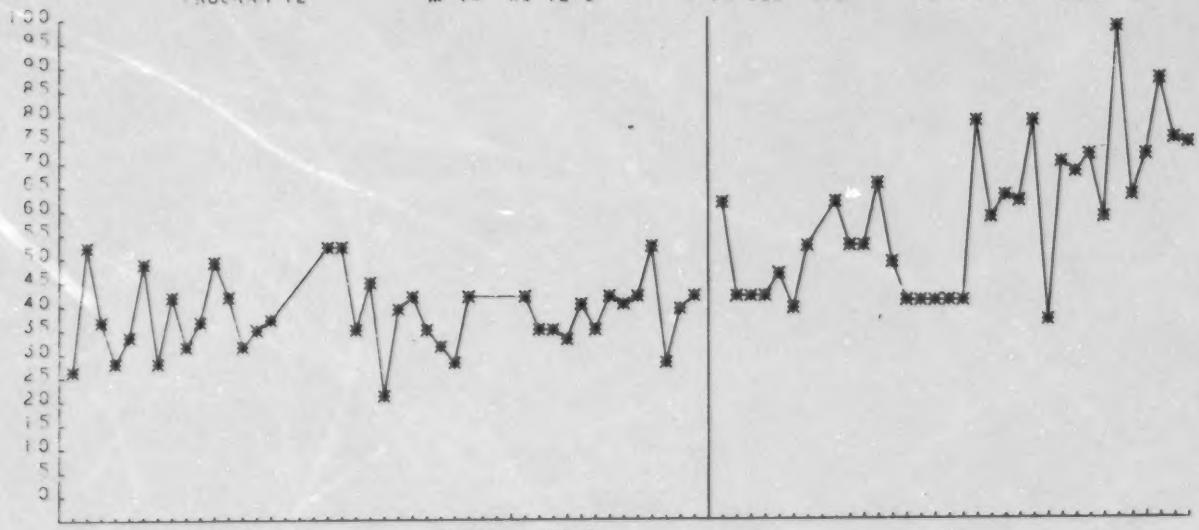
Program and Variable Identification: The program number is given below each abscissa, followed by the symbols for each variable and a number identifying the variables plotted. This is followed by a multiplier, required so that all variables can be plotted on the same axis ranging from 1 to 100. These are described below.

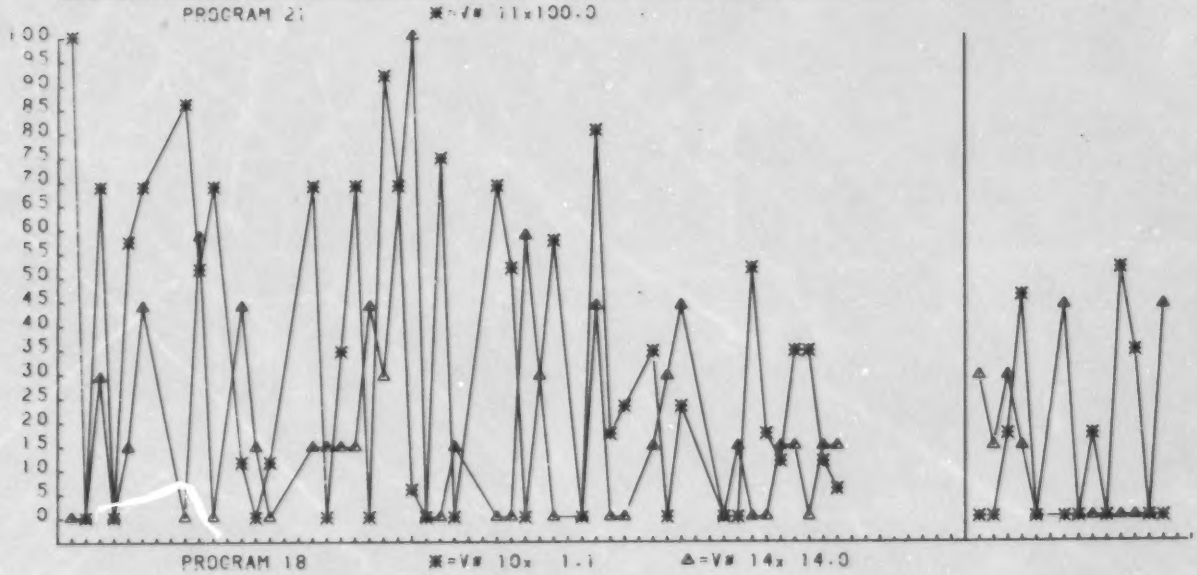
Program Number	Program Name	Variable Number	Variable Name
12	Whining	9	Number of whining episodes
		10	Average duration each "whine"
		11	Average intensity each "whine"
13	Fighting	9	Number of fighting episodes
		10	Average duration of each fight
		11	Average intensity of each fight
15	Tantrums	9	Number of tantrum episodes
		10	Average duration of each tantrum
		11	Average intensity of each tantrum
17	Interrupting	9	Number of interruption episodes
		10	Average duration of each interruption
		11	Average intensity of each interruption
18	Bedtime	10	Minutes late going to bed
		14	Number of times cried, yelled or out-of-bed
20	Room Neatness	11	Proportion of tasks completed
21	Following Directions	11	Proportion of directions followed

Treatment: The implementation of treatment in the time-lagged multiple-baseline design is always indicated by a vertical line on the graph.

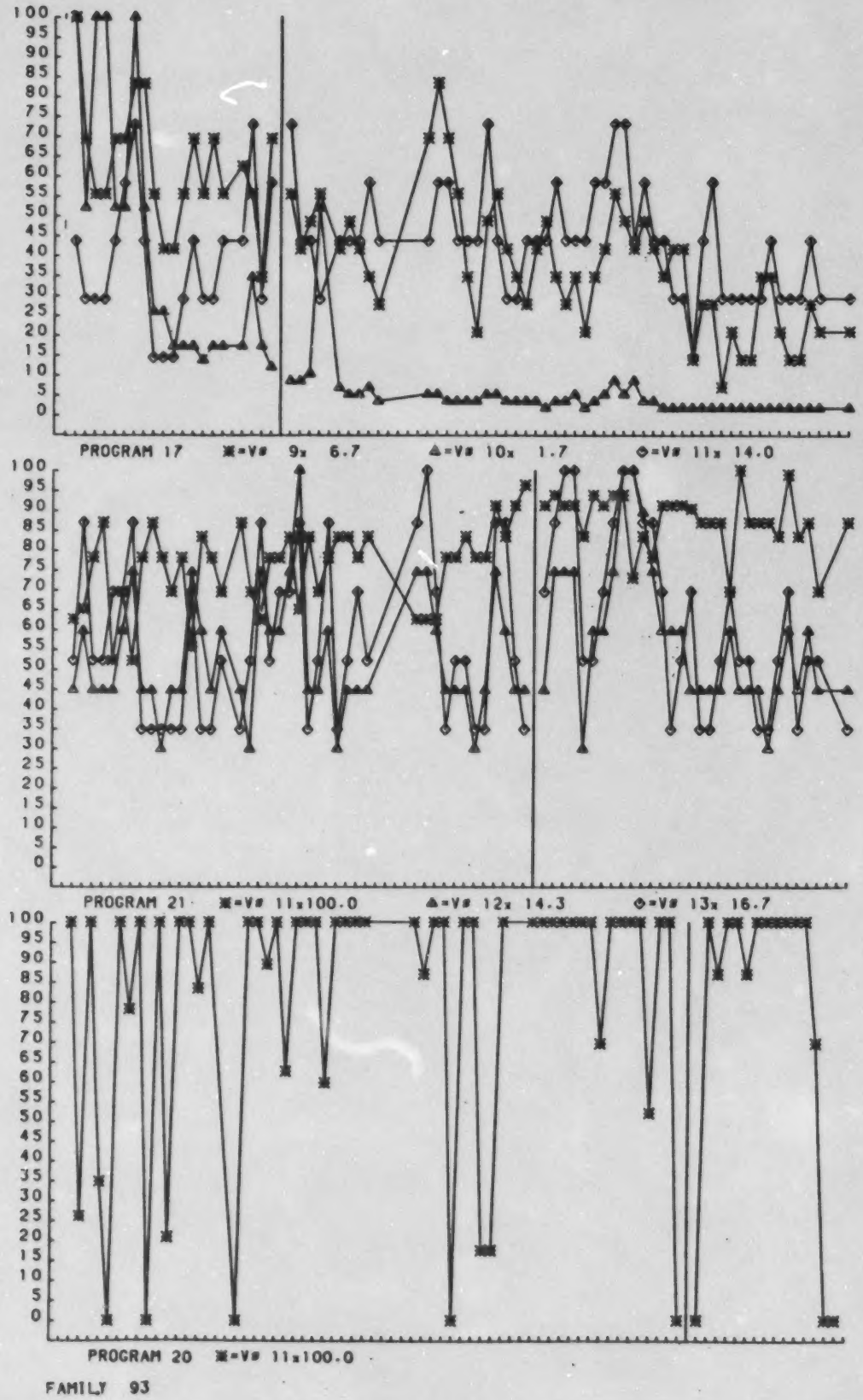
Variable Symbols: One to three variables are printed on each set of axes. For ease of reading, the variable plotted with an asterisk (*) is always the major dependent measure, and the one used in other reported analyses. It is always frequency, proportion completed or followed, or an equivalent, except in bedtime where it is minutes late.

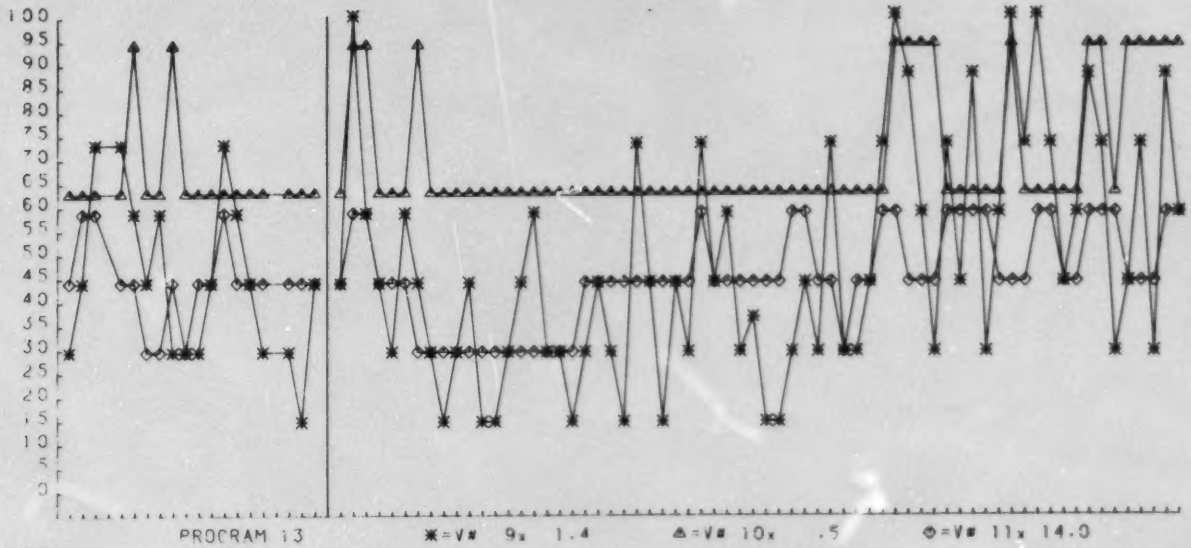

PROGRAM 12
*=V# 9x 12.0
△=V# 10x 1.0
◇=V# 11x 14.0

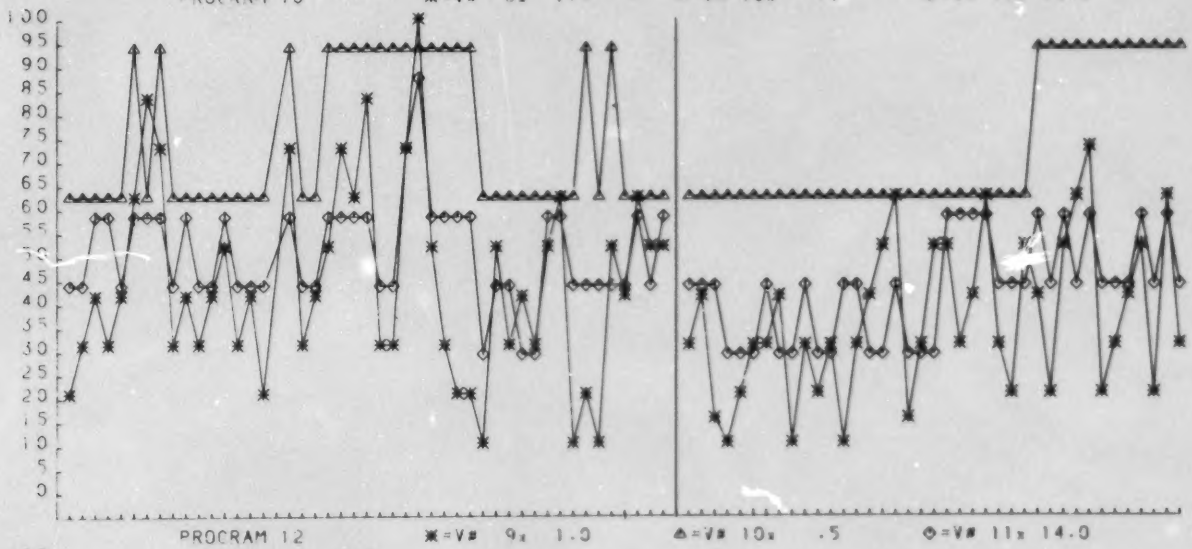

PROGRAM 21
*=V# 11x100.0

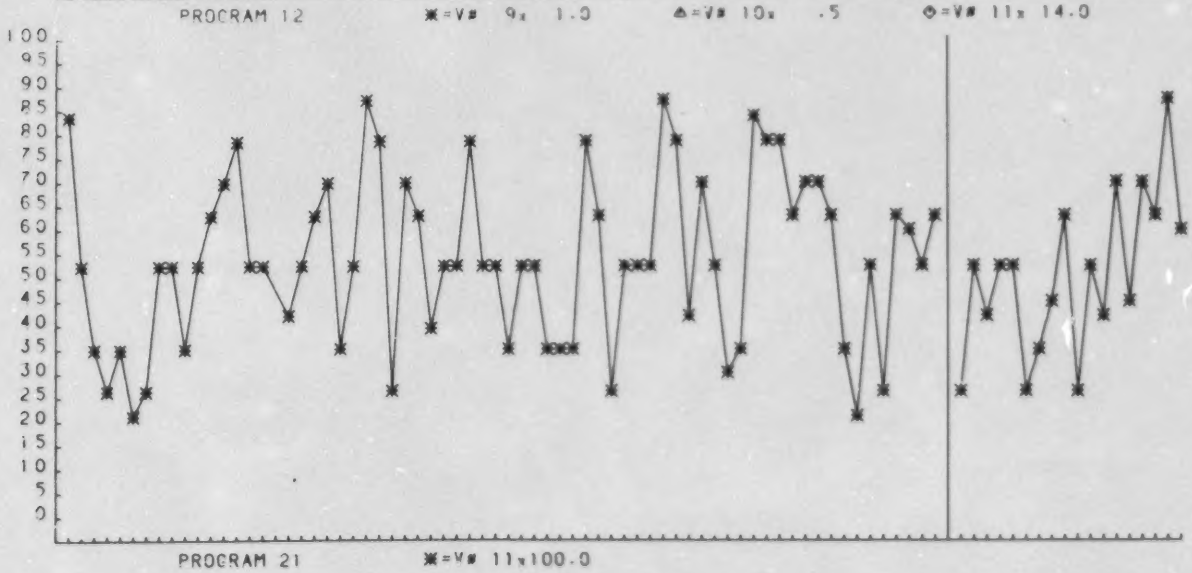

PROGRAM 18
*=V# 10x 1.1
△=V# 14x 14.0

FAMILY 91

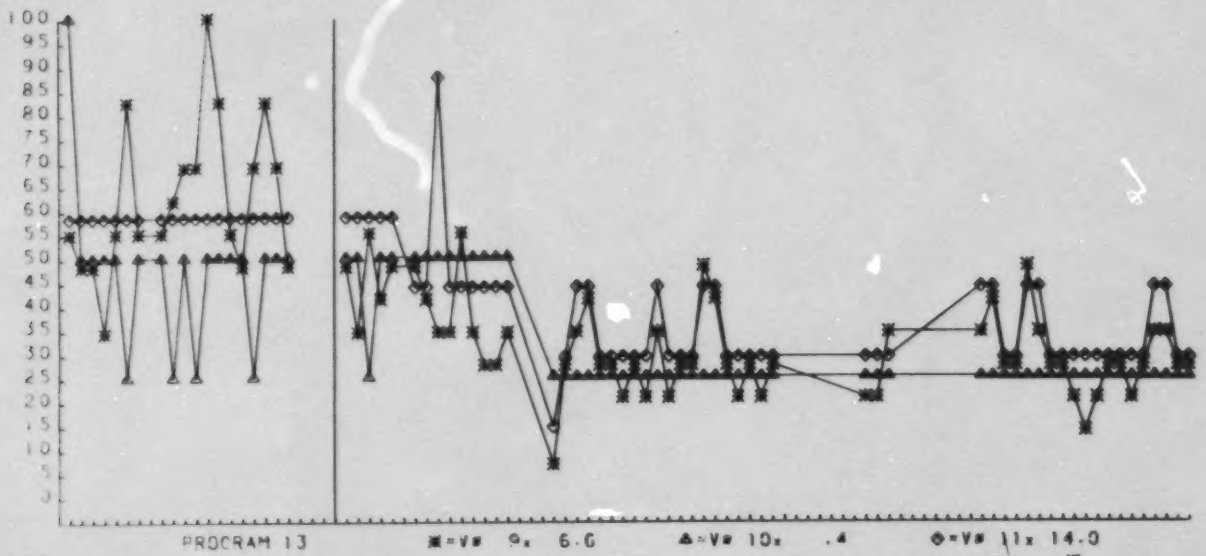

PROGRAM 17 *=V# 9x 6.7 △=V# 10x 1.7 ◇=V# 11x 14.0
PROGRAM 21 *=V# 11x100.0 △=V# 12x 14.3 ◇=V# 13x 16.7
PROGRAM 20 *=V# 11x100.0
FAMILY 93

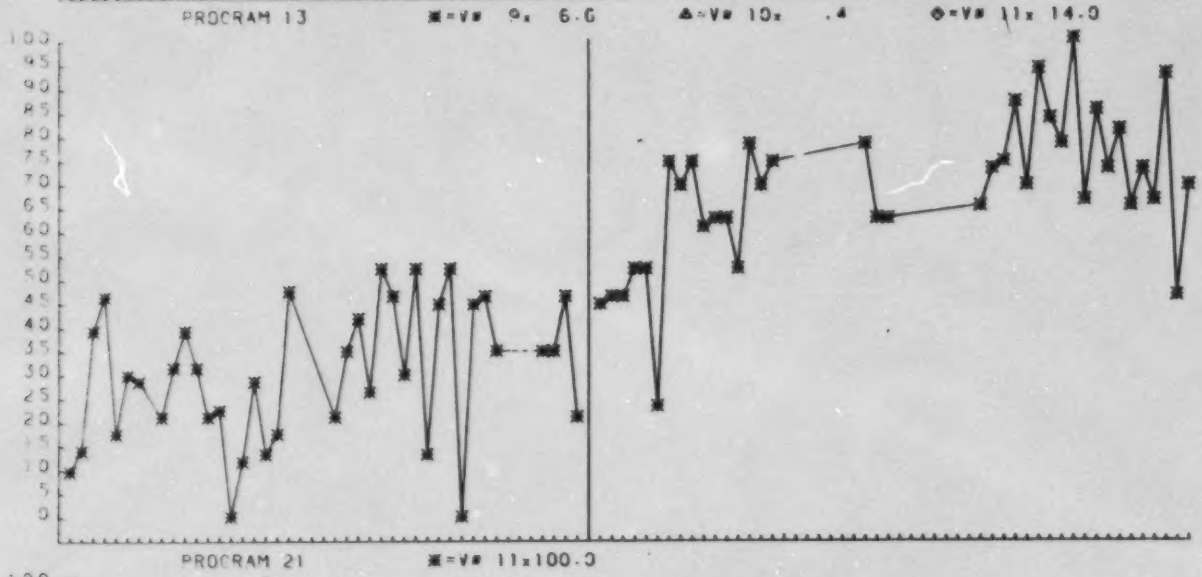

PROGRAM 13
*=V# 9x 1.4
△=V# 10x .5
◇=V# 11x 14.0

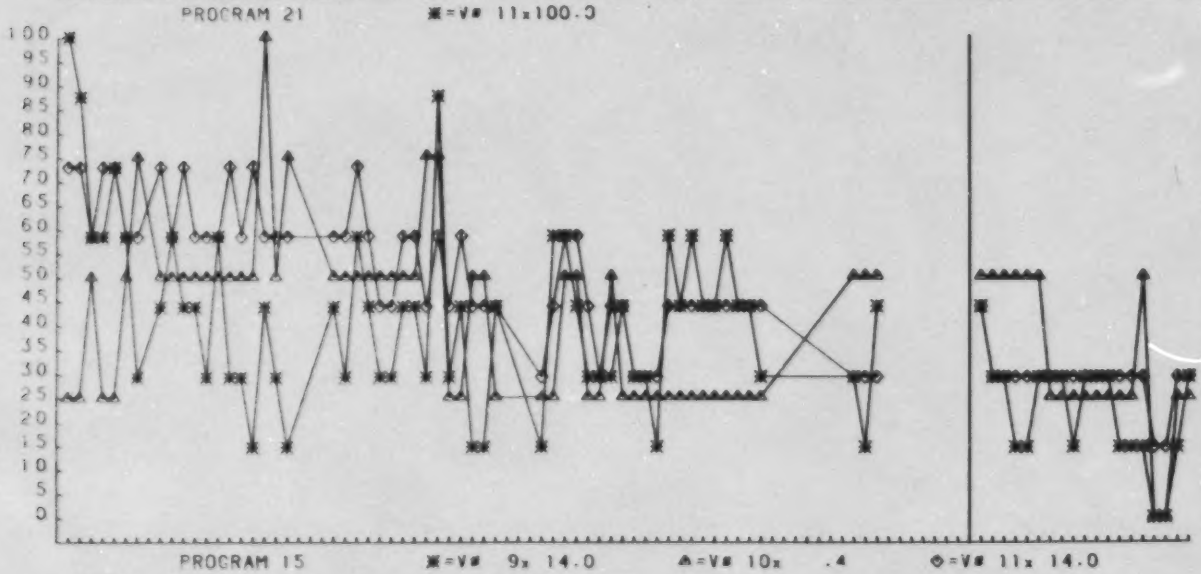

PROGRAM 12
*=V# 9x 1.0
△=V# 10x .5
◇=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0

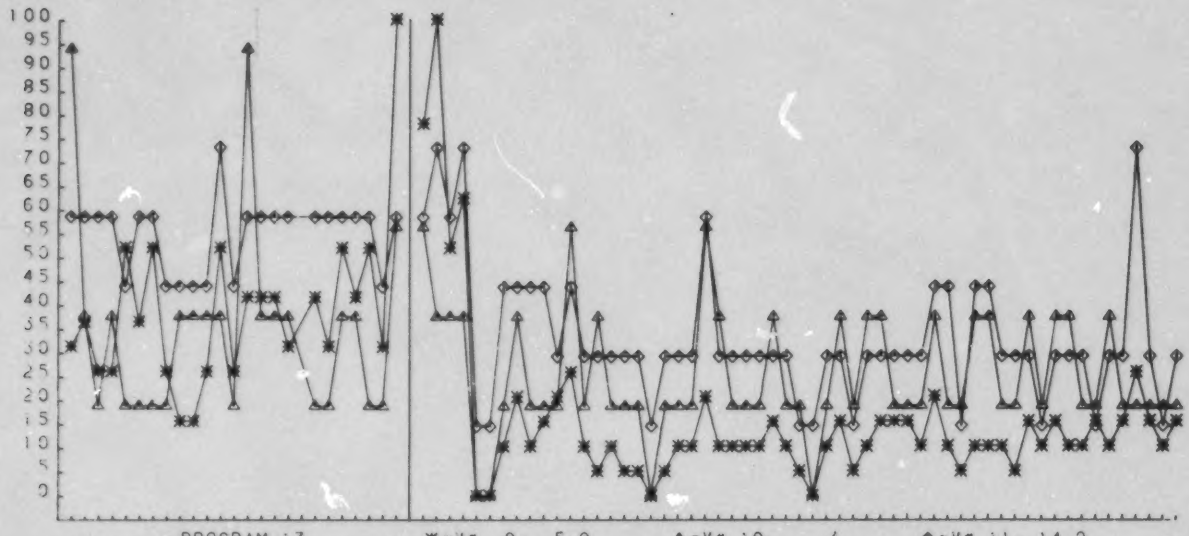
FAMILY 101

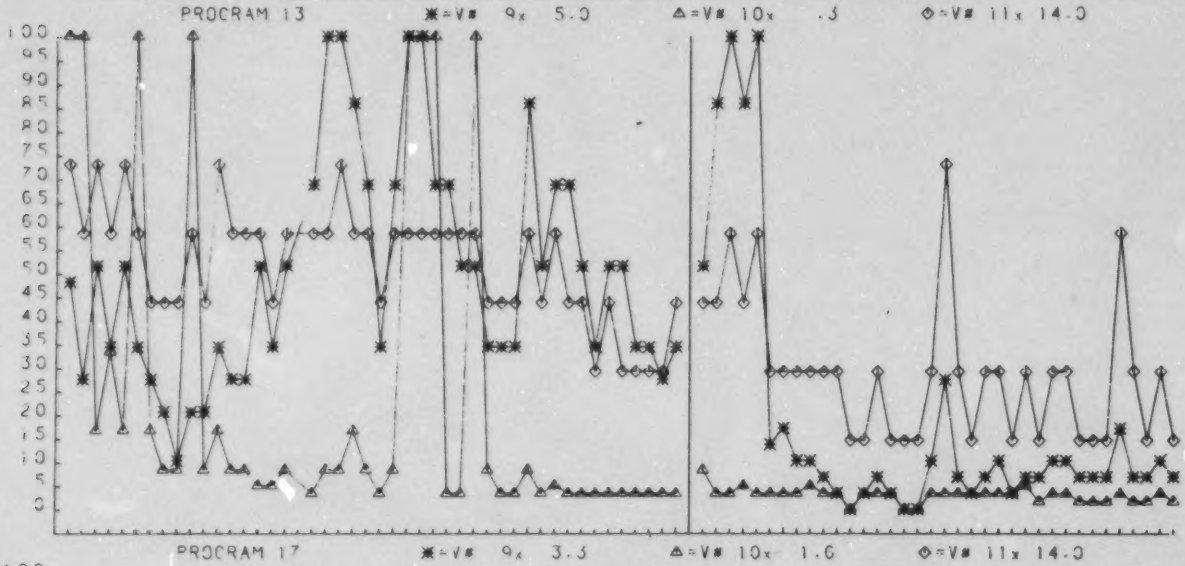

PROGRAM 13
*=V# 9x 6.6
▲=V# 10x .4
◆=V# 11x 14.0

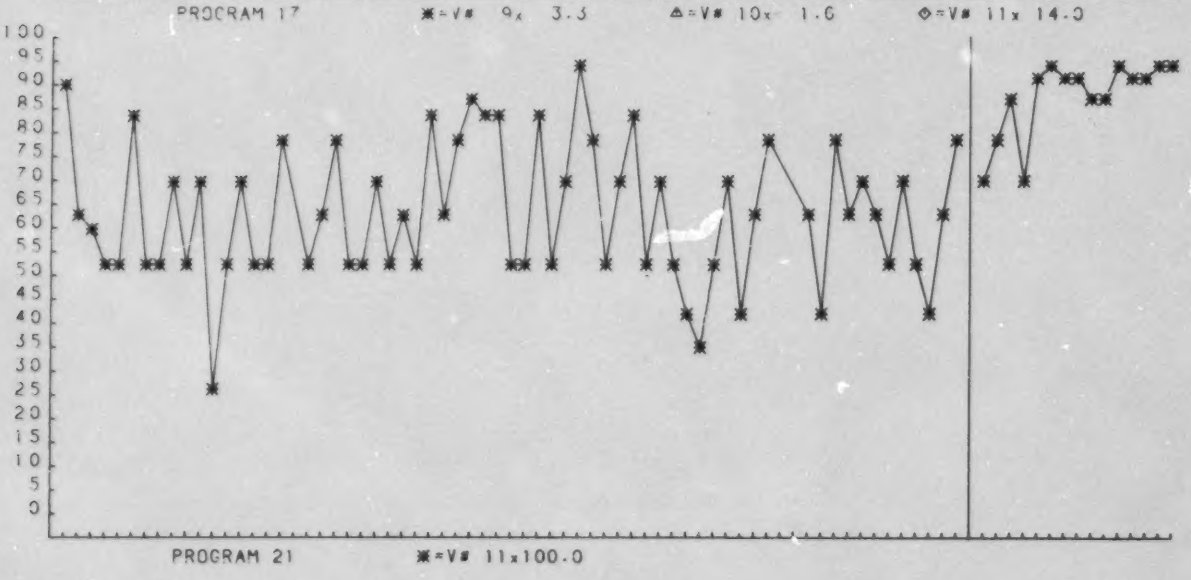

PROGRAM 21
*=V# 11x100.3


PROGRAM 15
*=V# 9x 14.0
▲=V# 10x .4
◆=V# 11x 14.0

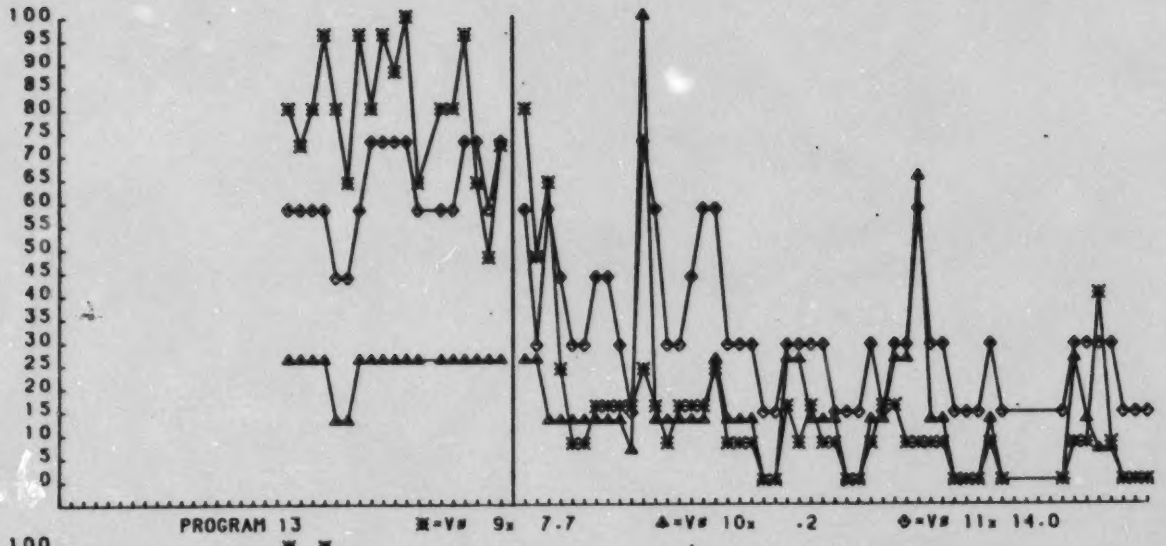
FAMILY 103

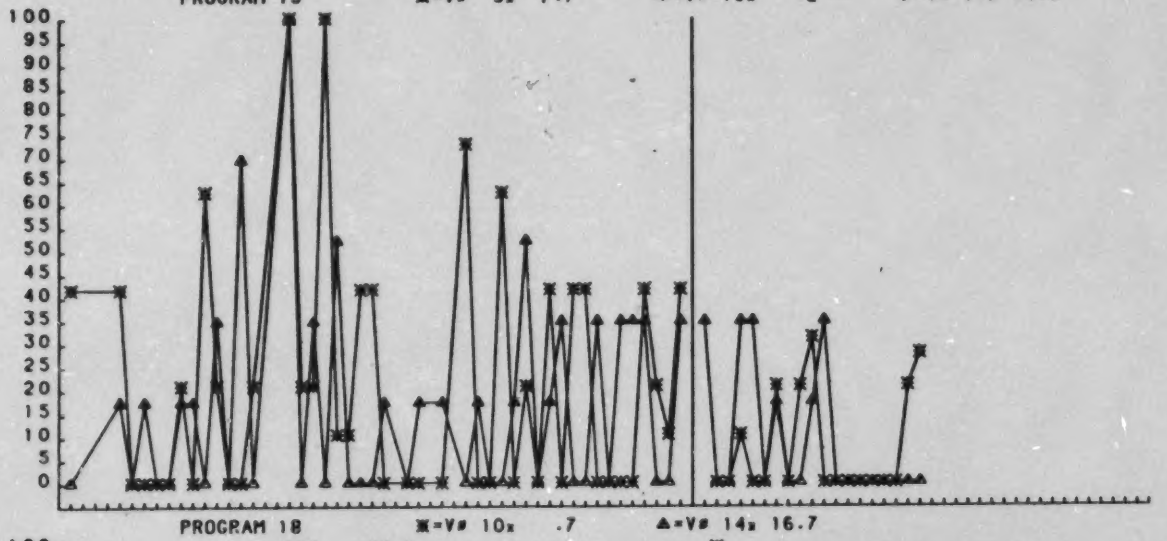

PROGRAM 13
*=V# 9x 5.0
△=V# 10x .3
◇=V# 11x 14.0

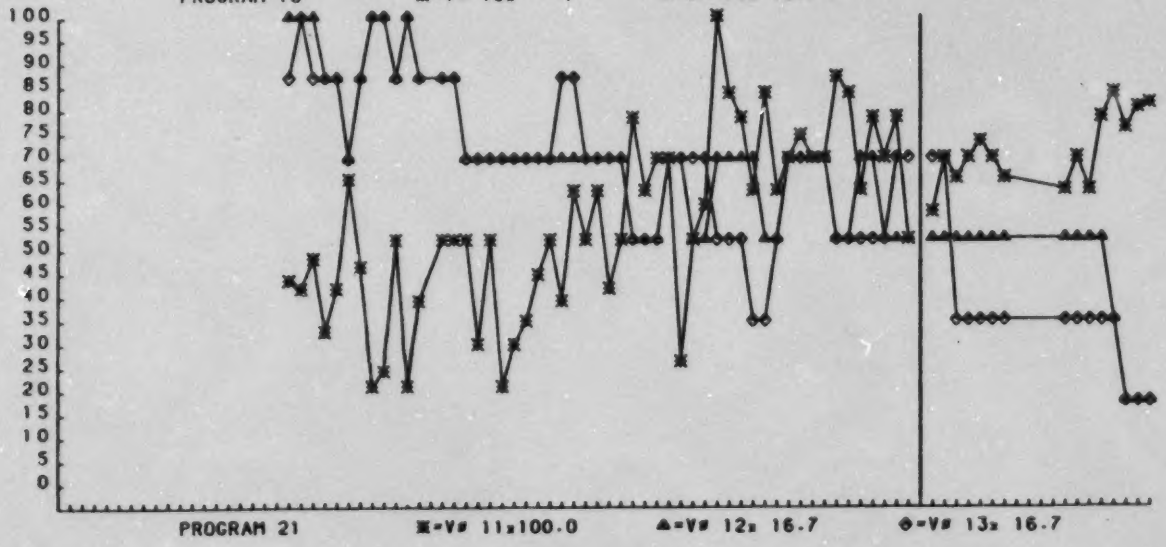

PROGRAM 17
*=V# 9x 3.3
△=V# 10x 1.0
◇=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0

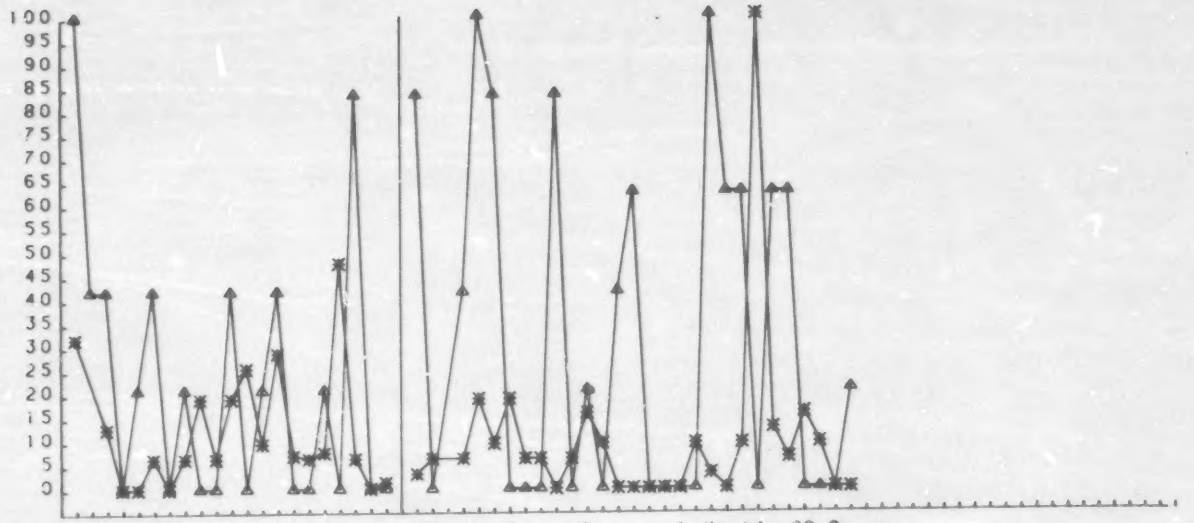
FAMILY 107

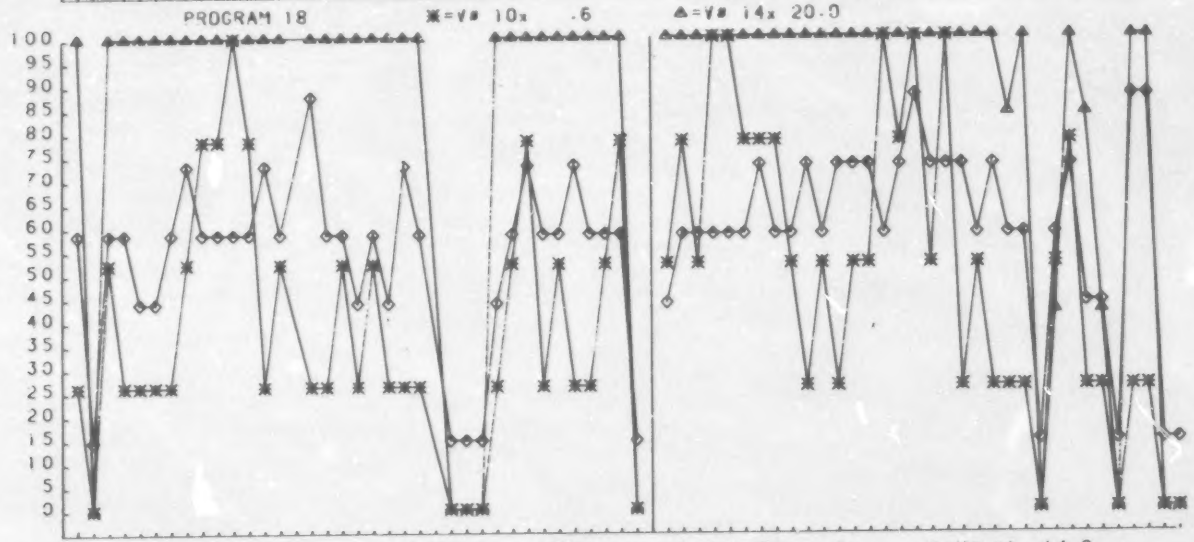

PROGRAM 13
*=V# 9x 7.7
▲=V# 10x .2
◆=V# 11x 14.0

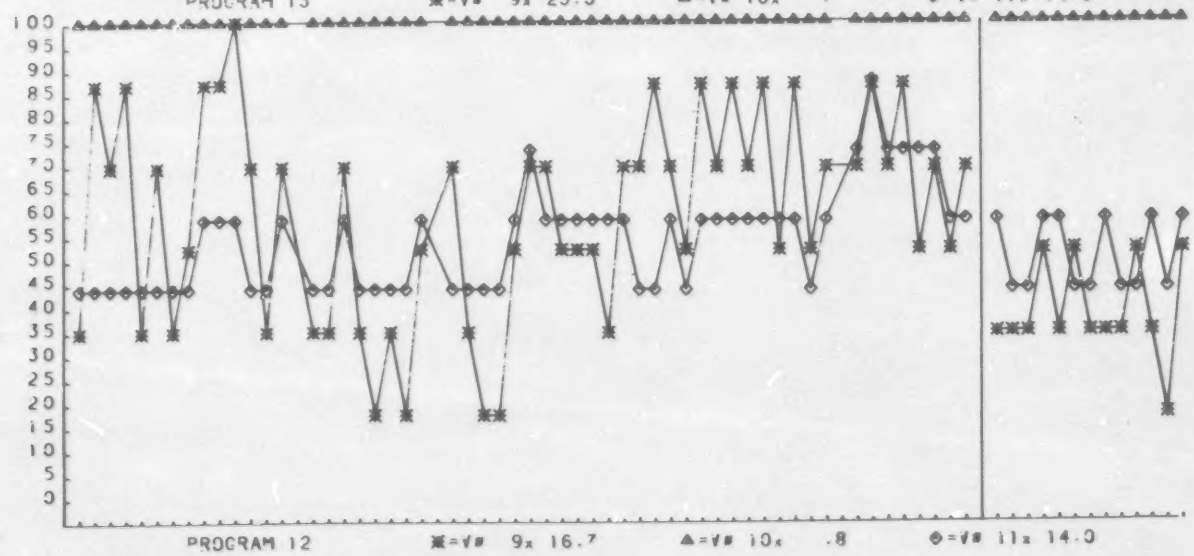

PROGRAM 18
*=V# 10x .7
▲=V# 14x 16.7


PROGRAM 21
*=V# 11x100.0
▲=V# 12x 16.7
◆=V# 13x 16.7

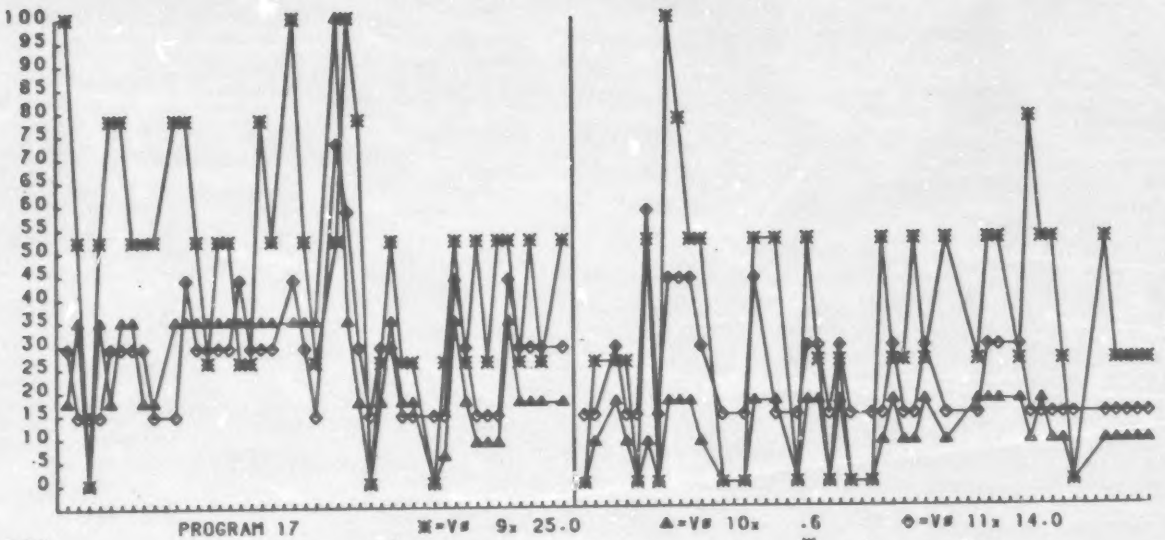
FAMILY 112

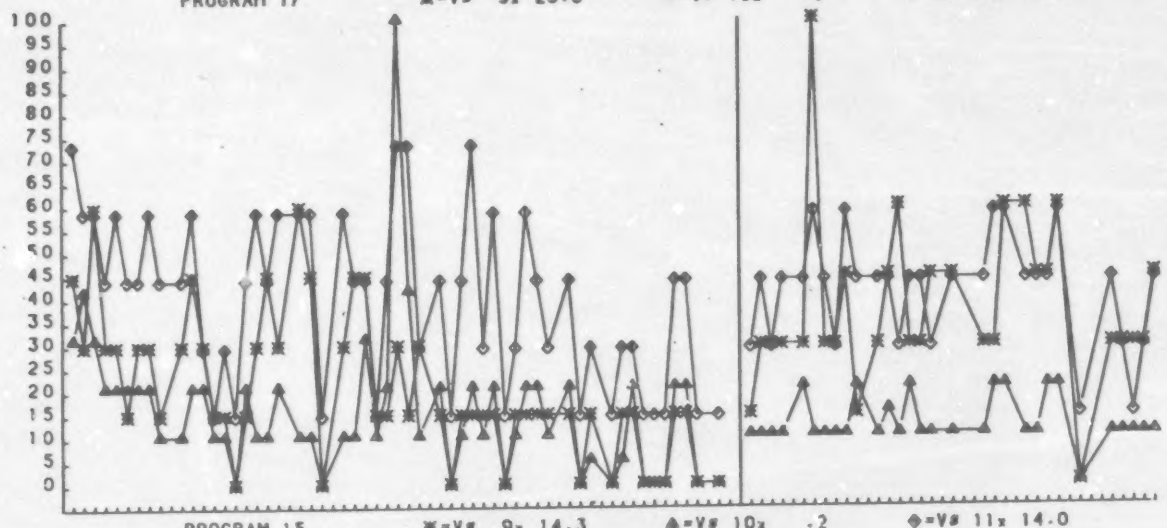

PROGRAM 18
*=V# 10x .6
▲=V# 14x 20.0

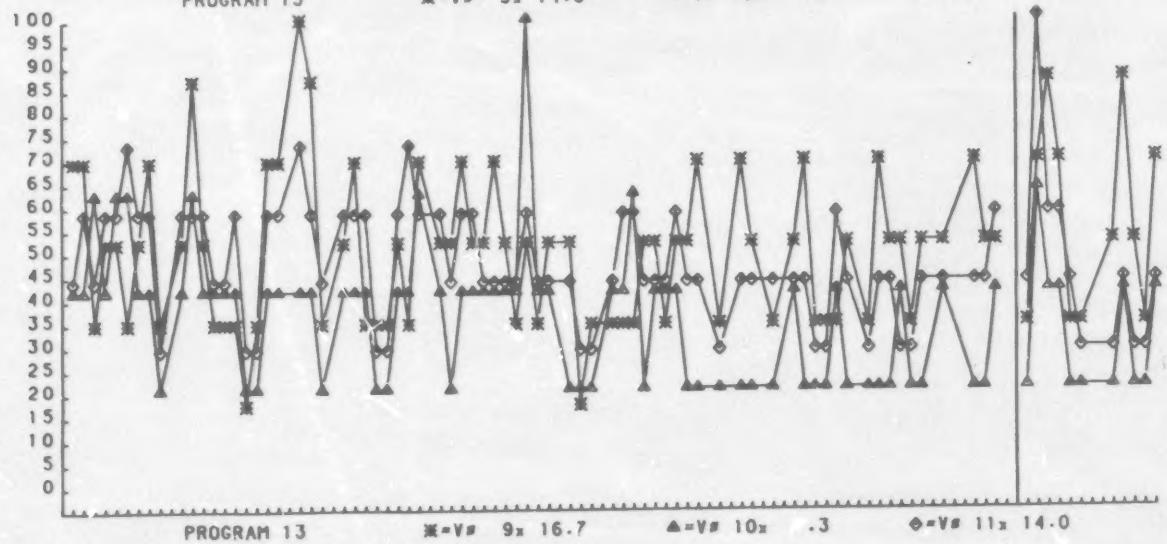

PROGRAM 15
*=V# 9x 25.0
▲=V# 10x .7
◇=V# 11x 14.0


PROGRAM 12
*=V# 9x 16.7
▲=V# 10x .8
◇=V# 11x 14.0

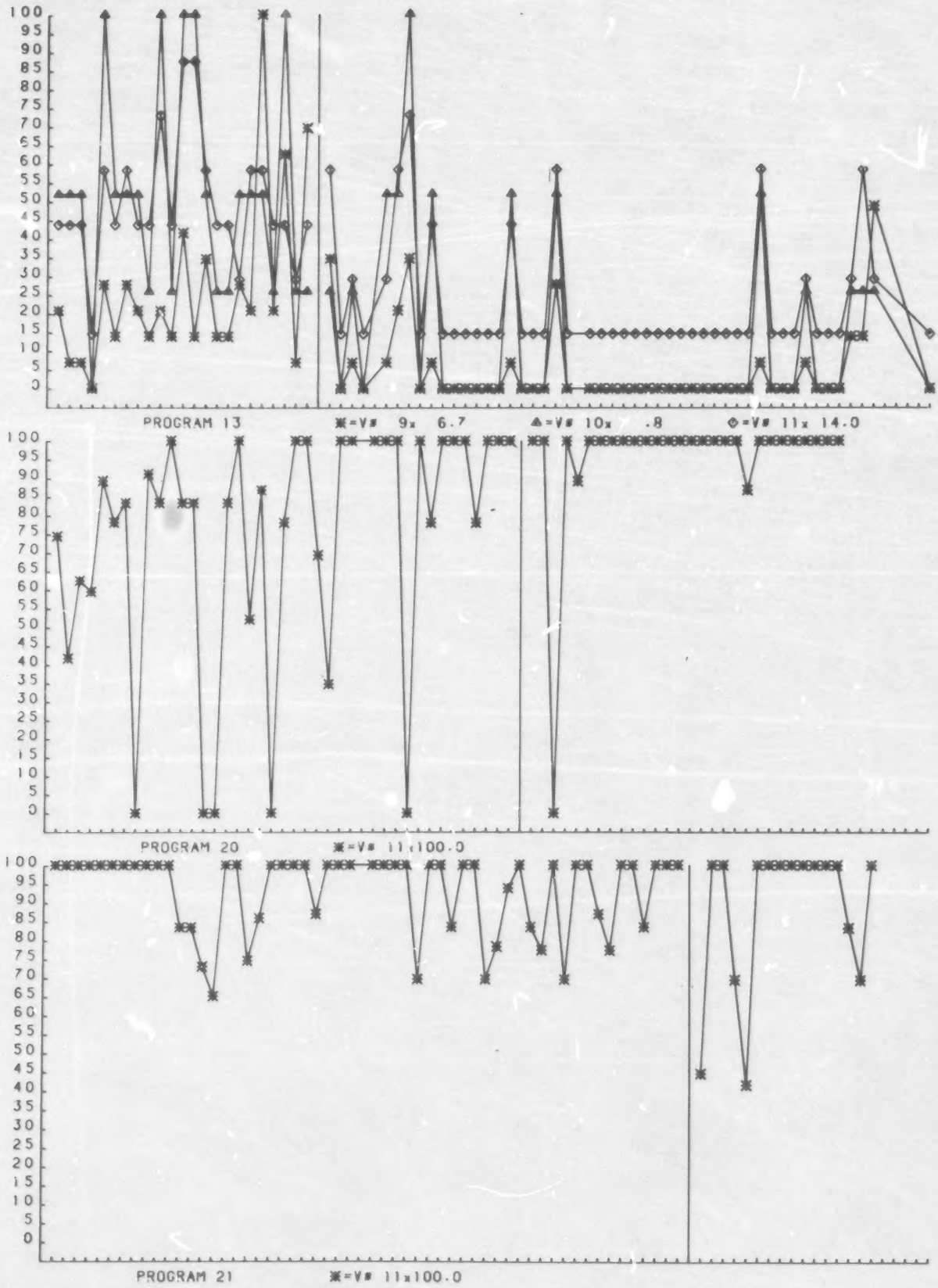
FAMILY 115


PROGRAM 17
*=V# 9x 25.0
▲=V# 10x .6
◆=V# 11x 14.0

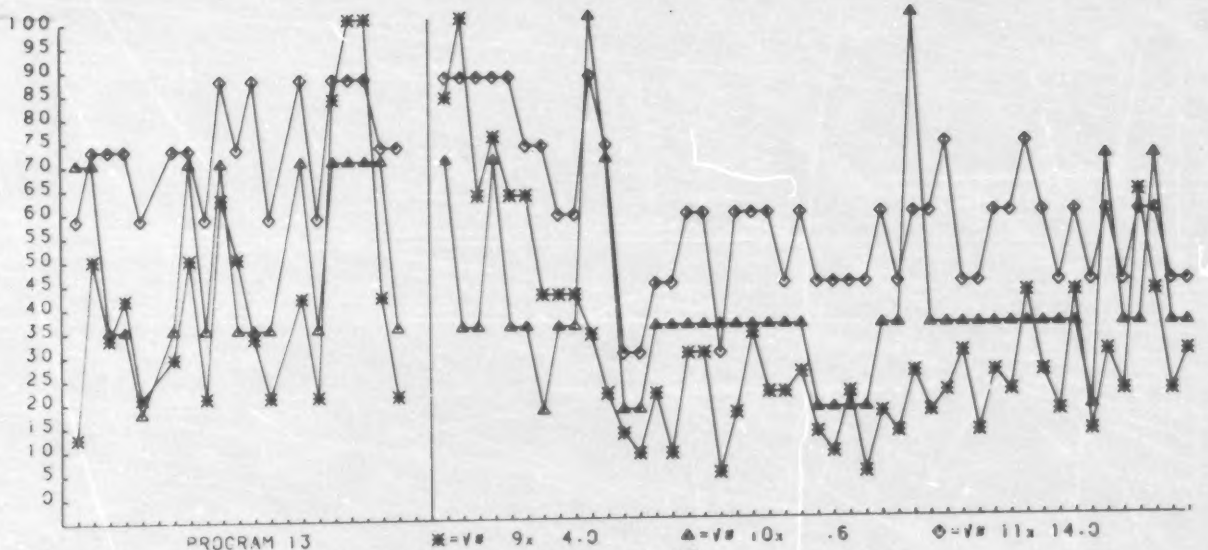

PROGRAM 15
*=V# 9x 14.3
▲=V# 10x .2
◆=V# 11x 14.0

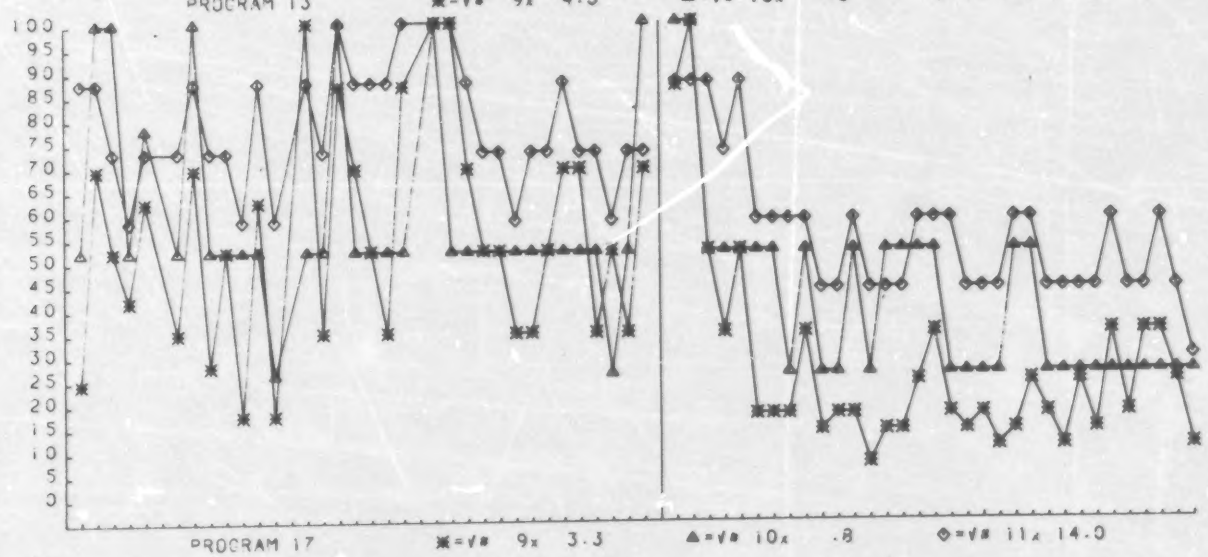

PROGRAM 13
*=V# 9x 16.7
▲=V# 10x .3
◆=V# 11x 14.0

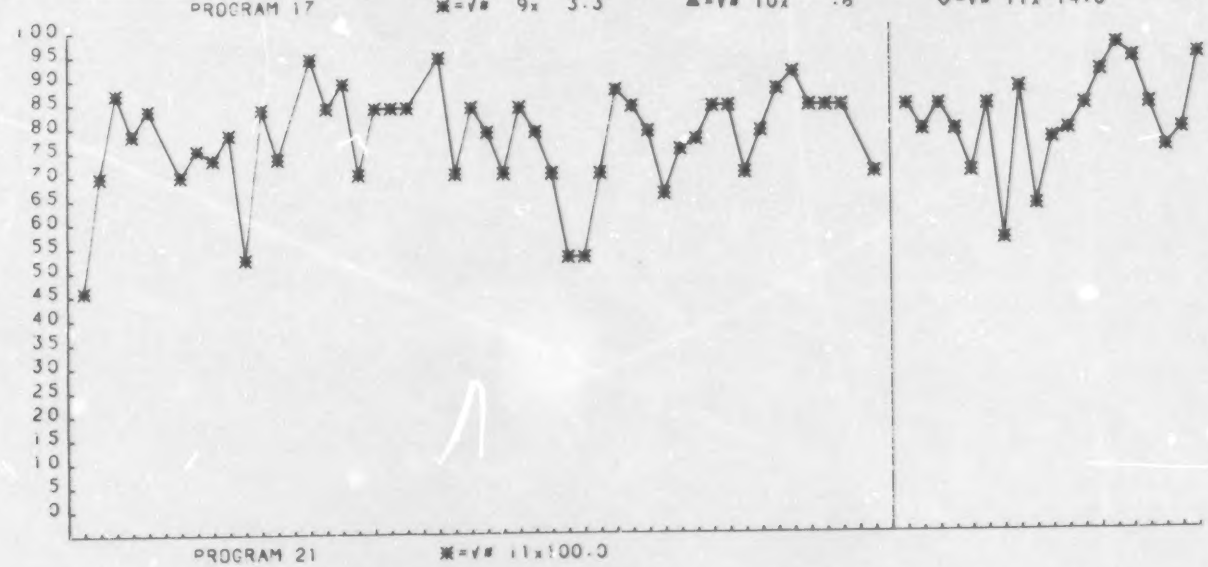
FAMILY 122


PROGRAM 13
*=V# 9x 6.7
▲=V# 10x .8
◇=V# 11x 14.0
PROGRAM 20
*=V# 11x100.0
PROGRAM 21
*=V# 11x100.0

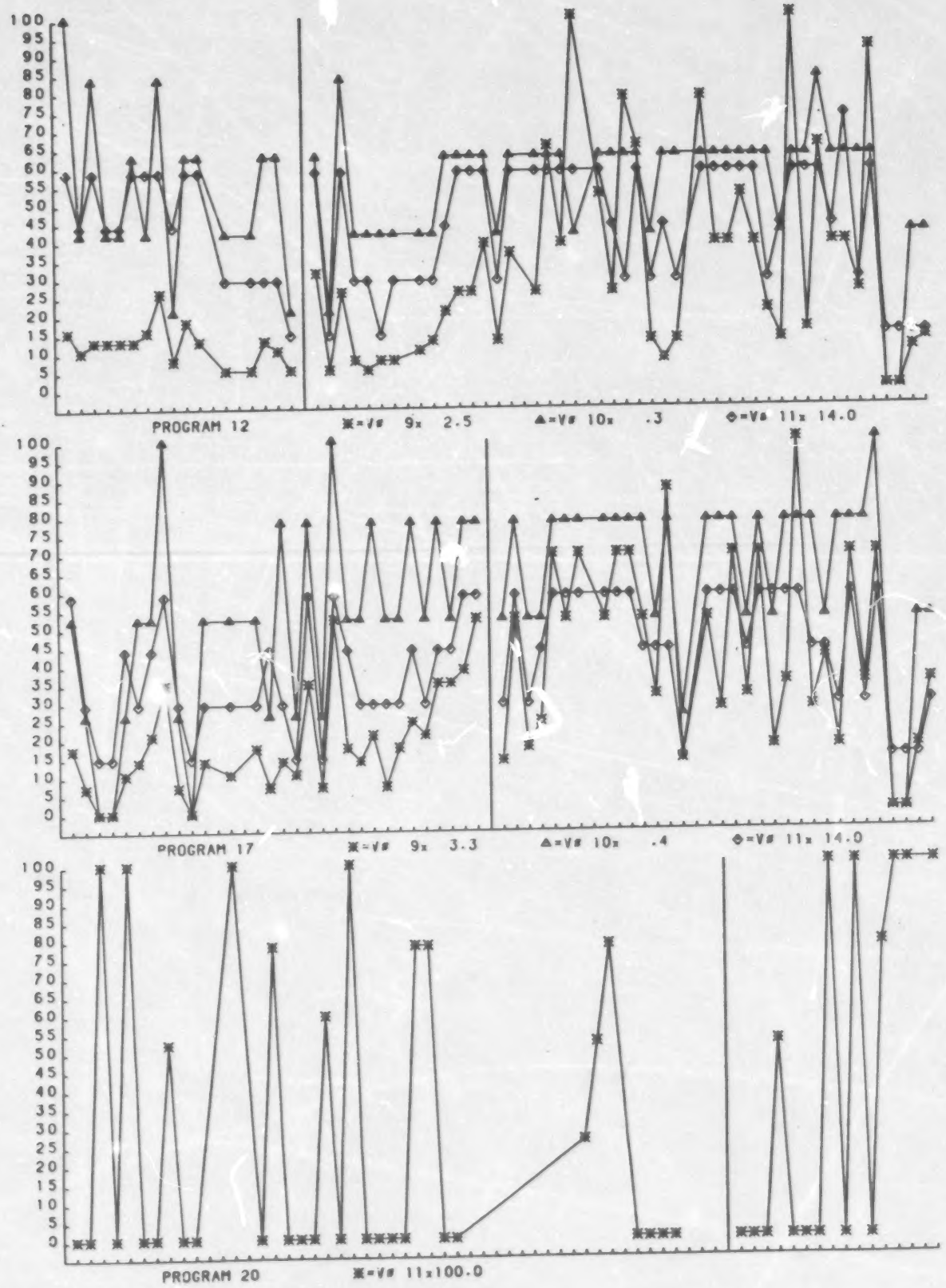
FAMILY 130


100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
PROGRAM 13
*=V# 9x 4.0
▲=V# 10x .6
◇=V# 11x 14.0

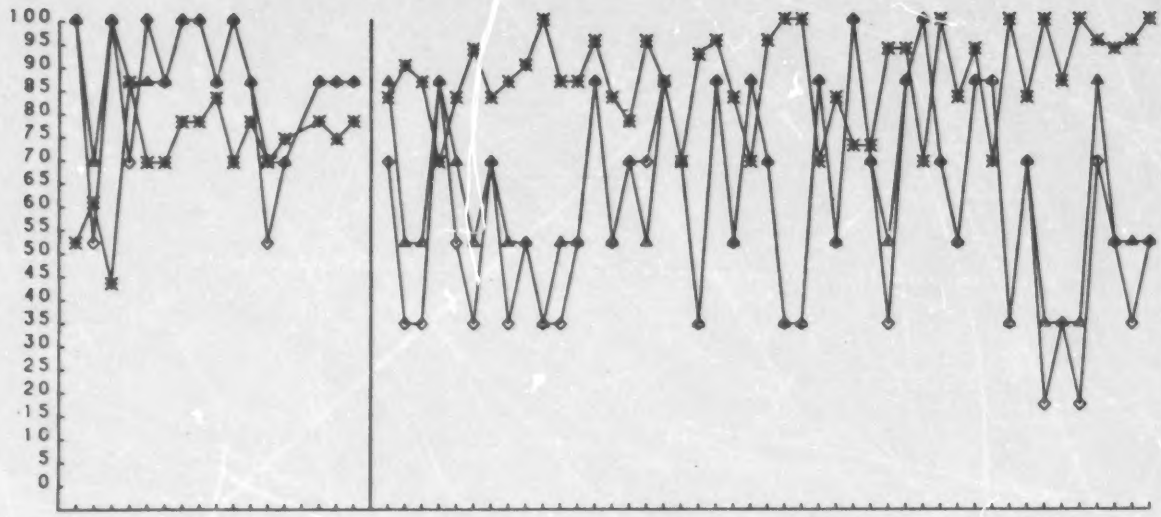

100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
PROGRAM 17
*=V# 9x 3.3
▲=V# 10x .8
◇=V# 11x 14.0

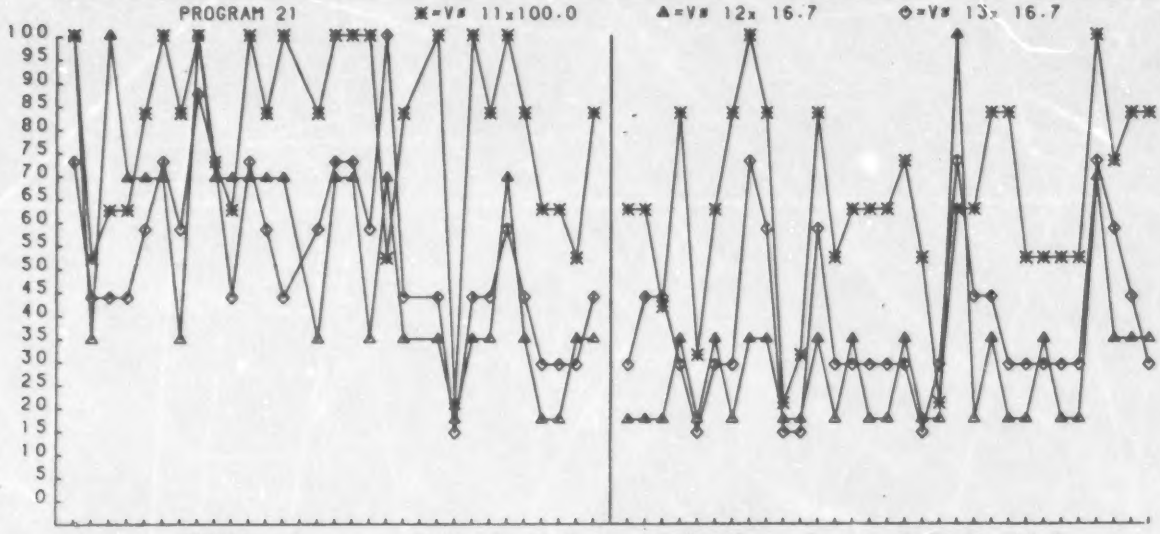

100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
PROGRAM 21
*=V# 11x100.0

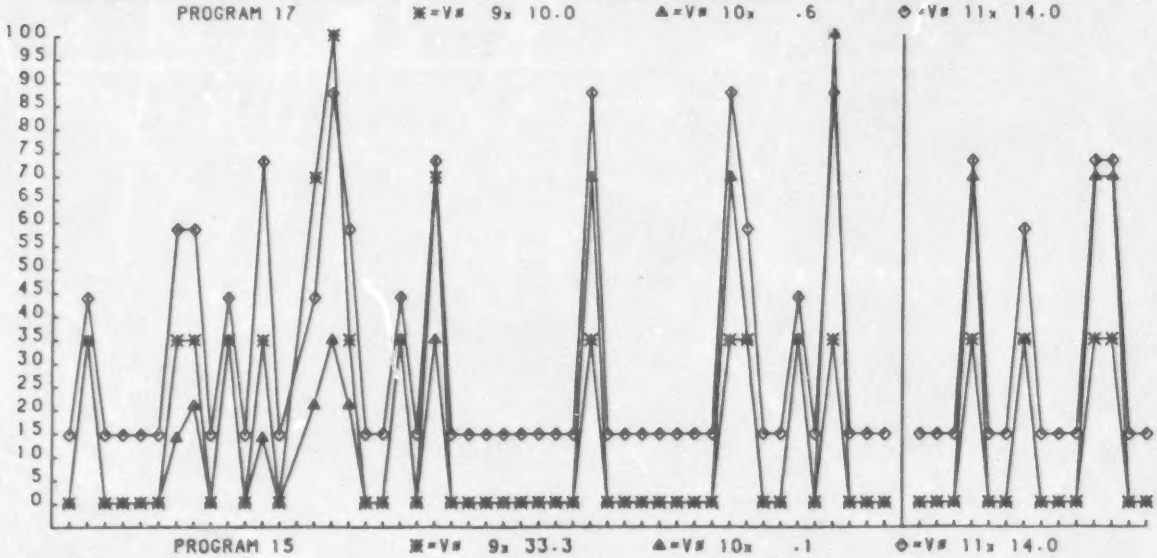
FAMILY 131


PROGRAM 12
*=V# 9x 2.5
▲=V# 10x .3
◇=V# 11x 14.0
PROGRAM 17
*=V# 9x 3.3
▲=V# 10x .4
◇=V# 11x 14.0
PROGRAM 20
*=V# 11x100.0

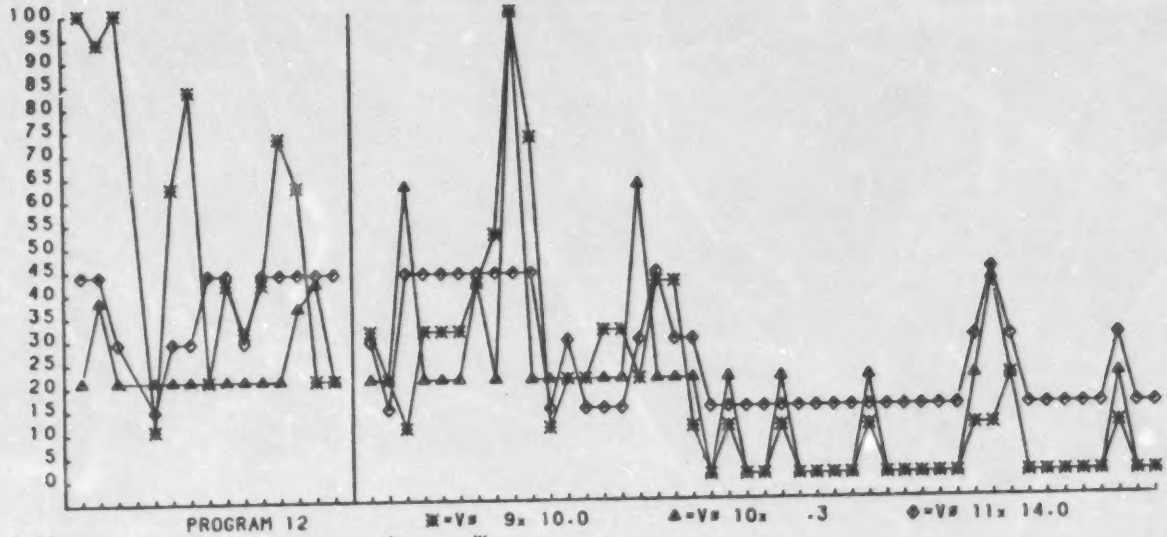
FAMILY 137

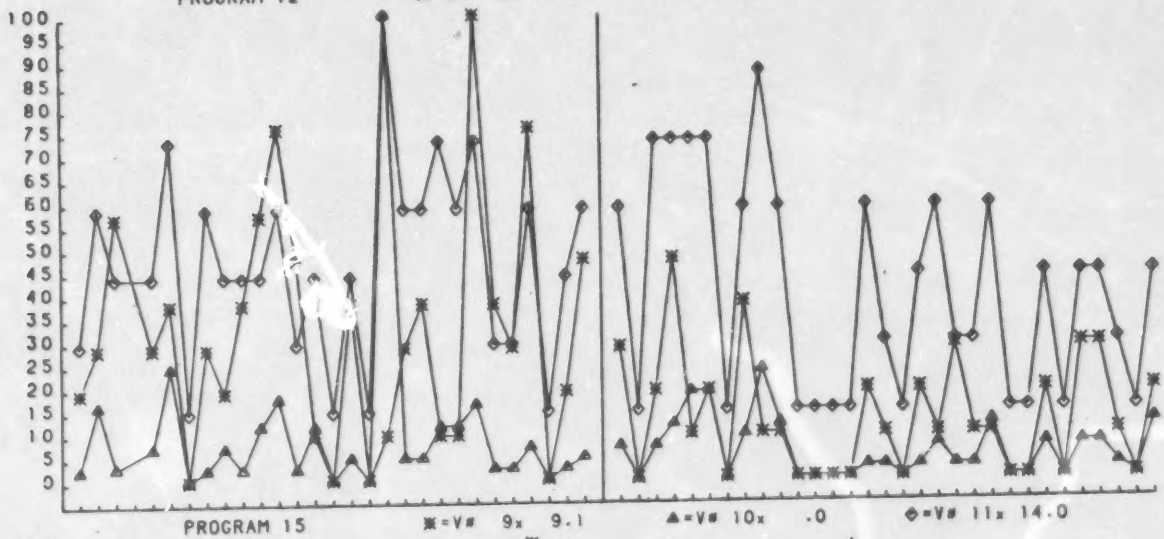

PROGRAM 21
*=V# 11x100.0
▲=V# 12x 16.7
◇=V# 13x 16.7

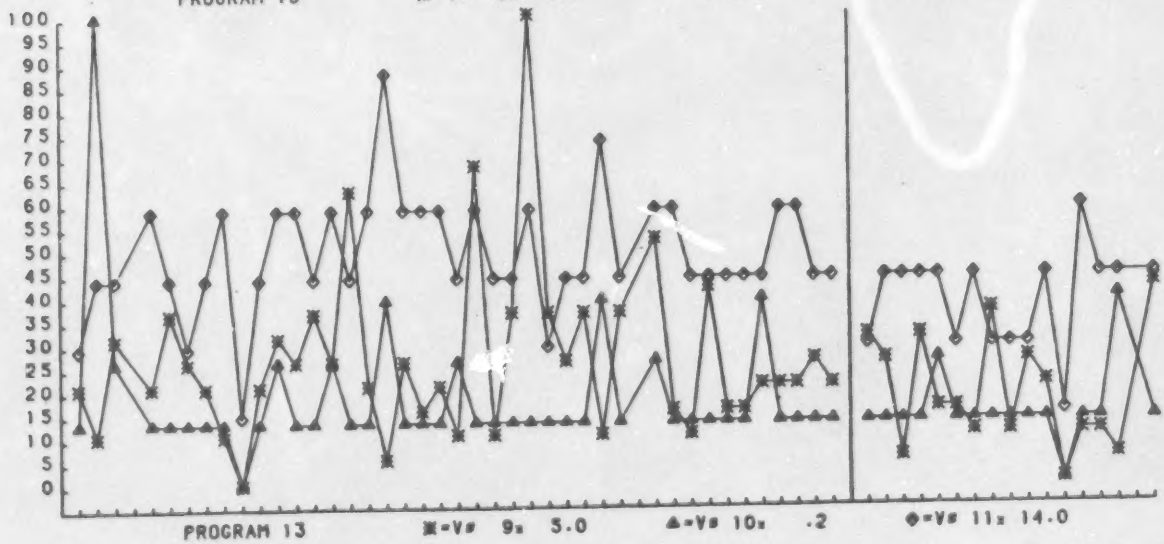

PROGRAM 17
*=V# 9x 10.0
▲=V# 10x .6
◇=V# 11x 14.0


PROGRAM 15
*=V# 9x 33.3
▲=V# 10x .1
◇=V# 11x 14.0

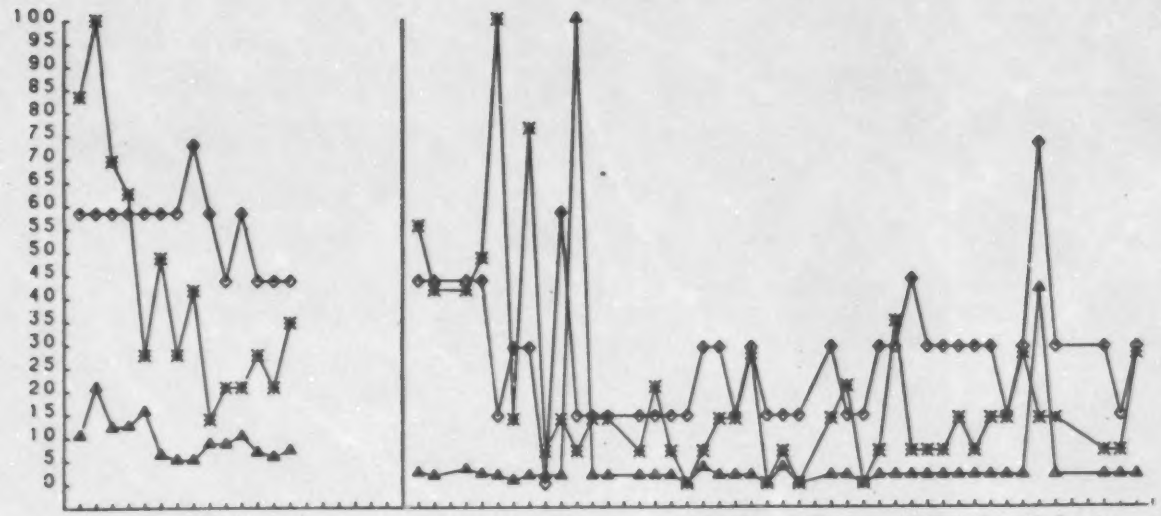
FAMILY 139

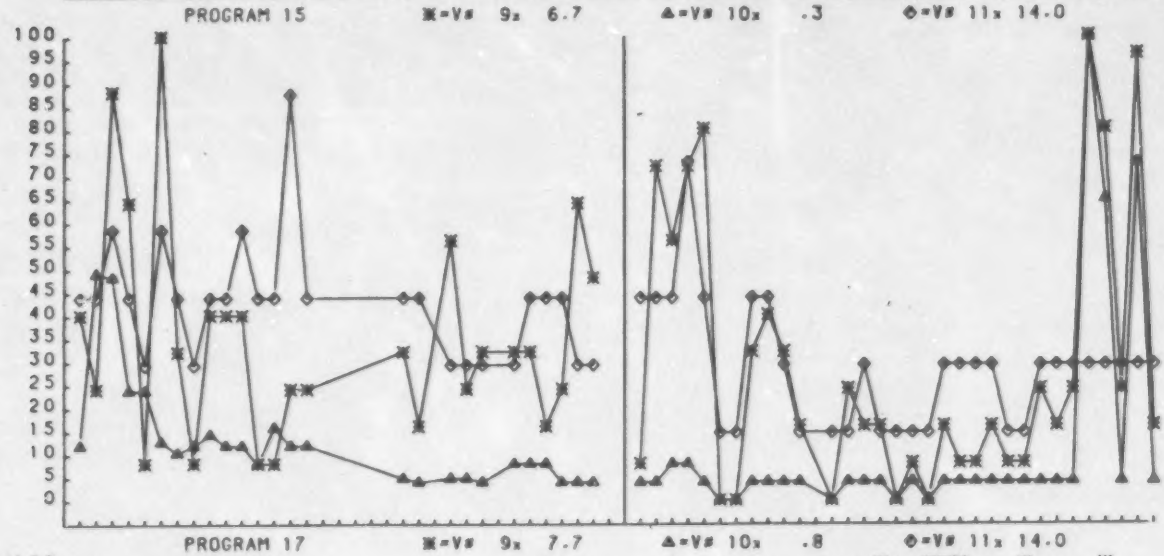

PROGRAM 12
✱=V# 9x 10.0
▲=V# 10x .3
◆=V# 11x 14.0

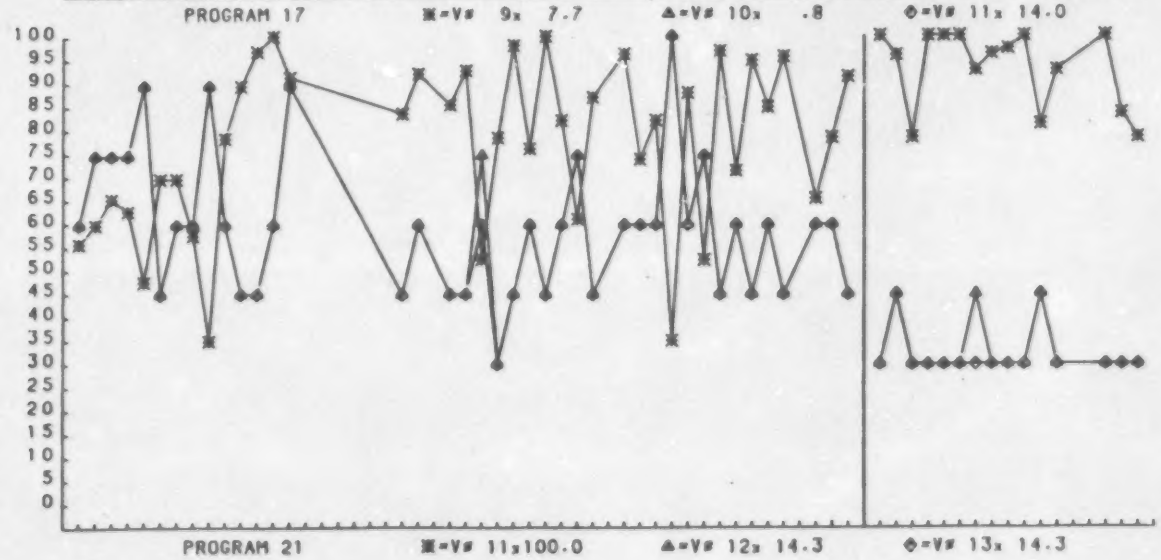

PROGRAM 15
✱=V# 9x 9.1
▲=V# 10x .0
◆=V# 11x 14.0


PROGRAM 13
✱=V# 9x 5.0
▲=V# 10x .2
◆=V# 11x 14.0

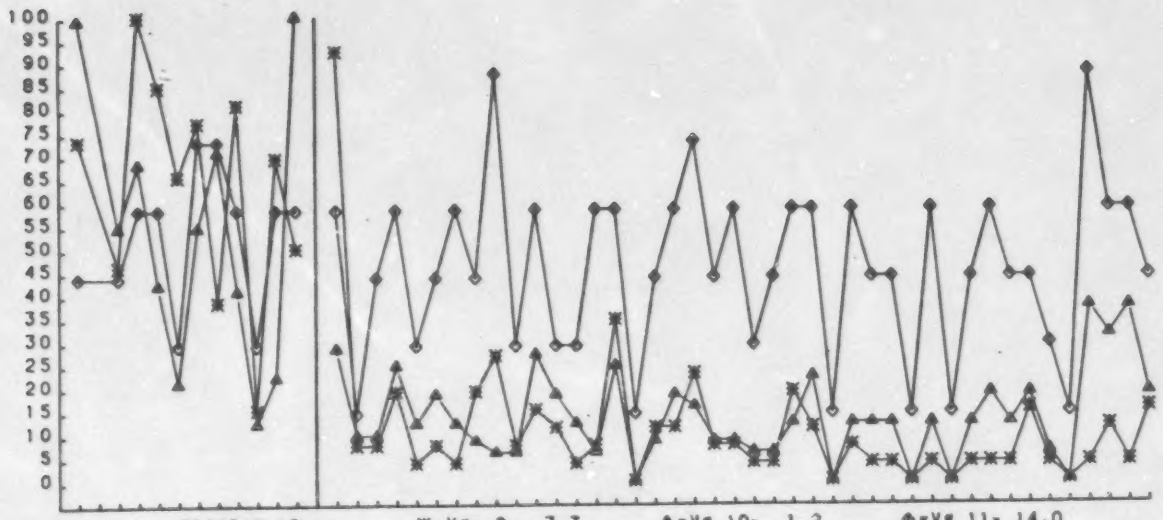
FAMILY 142

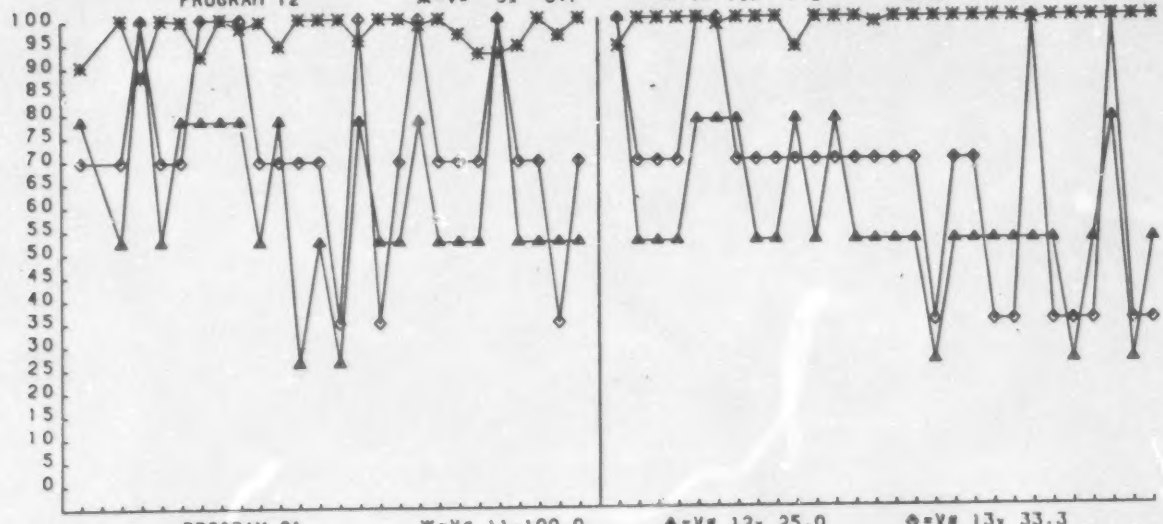

PROGRAM 15
✱=V# 9x 6.7
▲=V# 10x .3
◇=V# 11x 14.0

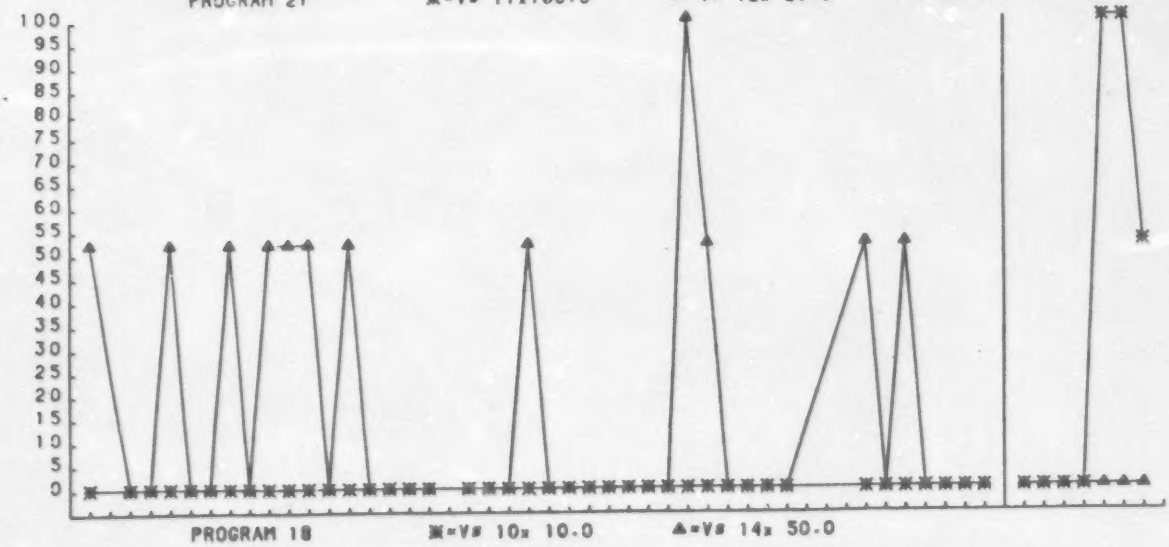

PROGRAM 17
✱=V# 9x 7.7
▲=V# 10x .8
◇=V# 11x 14.0


PROGRAM 21
✱=V# 11x100.0
▲=V# 12x 14.3
◇=V# 13x 14.3

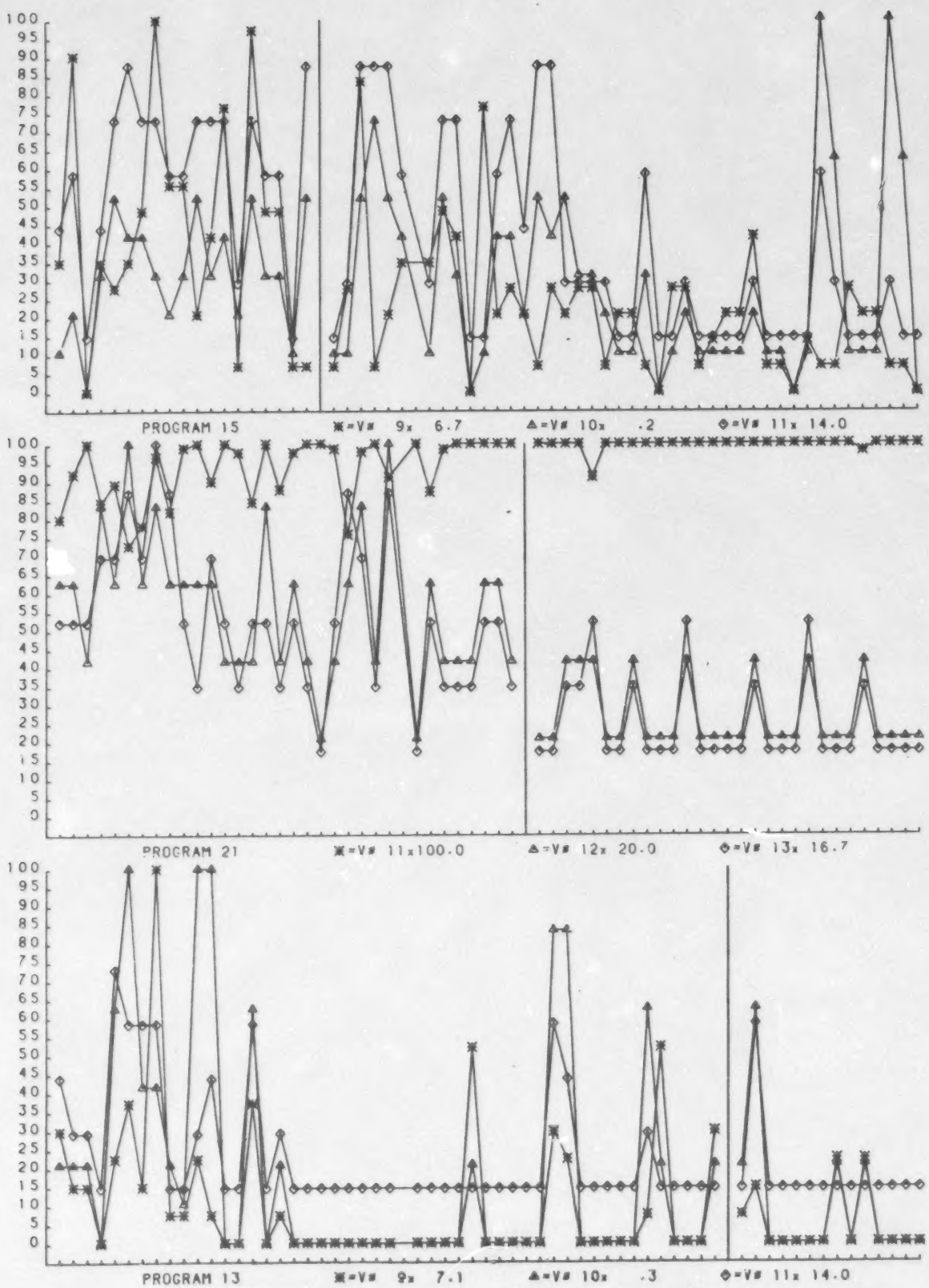
FAMILY 148


PROGRAM 12
*=V# 9x 3.7
▲=V# 10x 1.2
◆=V# 11x 14.0

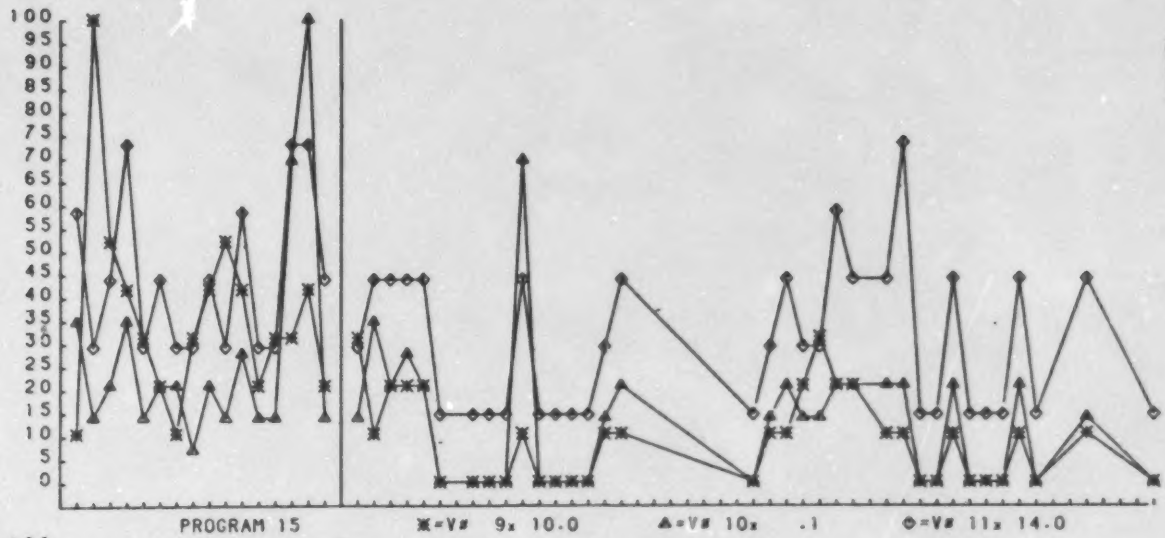

PROGRAM 21
*=V# 11x100.0
▲=V# 12x 25.0
◆=V# 13x 33.3

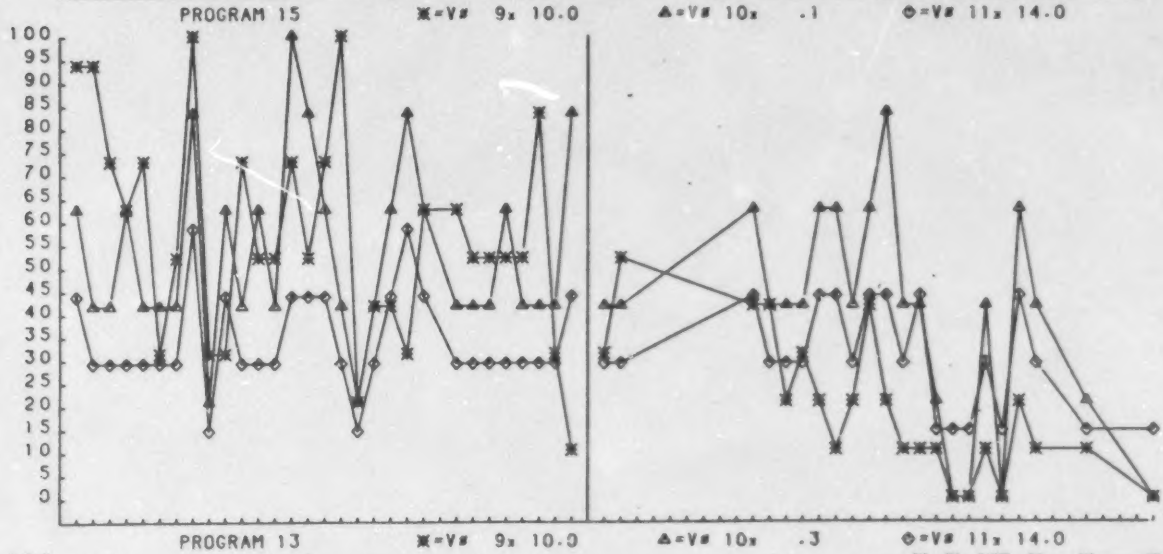

PROGRAM 18
*=V# 10x 10.0
▲=V# 14x 50.0

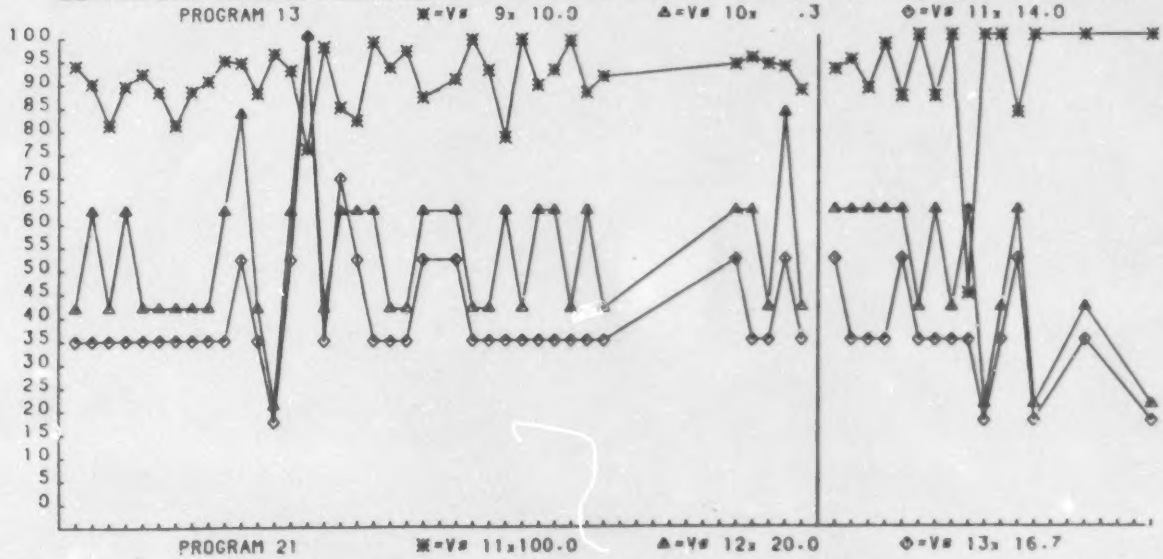
FAMILY 150


PROGRAM 15
✱=V# 9x 6.7
△=V# 10x .2
◇=V# 11x 14.0
PROGRAM 21
✱=V# 11x100.0
△=V# 12x 20.0
◇=V# 13x 16.7
PROGRAM 13
✱=V# 9x 7.1
△=V# 10x .3
◇=V# 11x 14.0

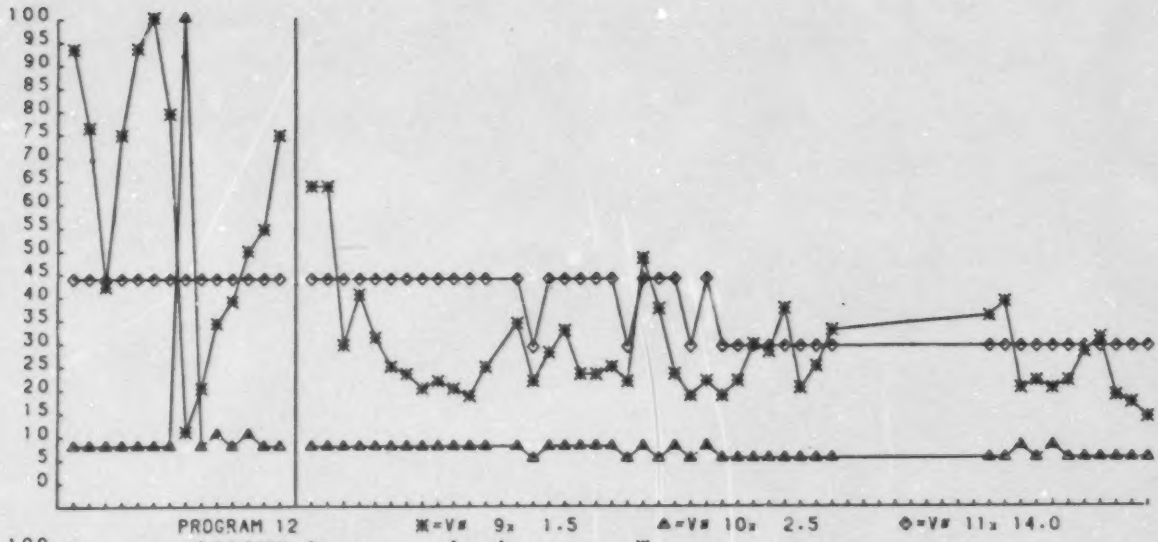
FAMILY 154

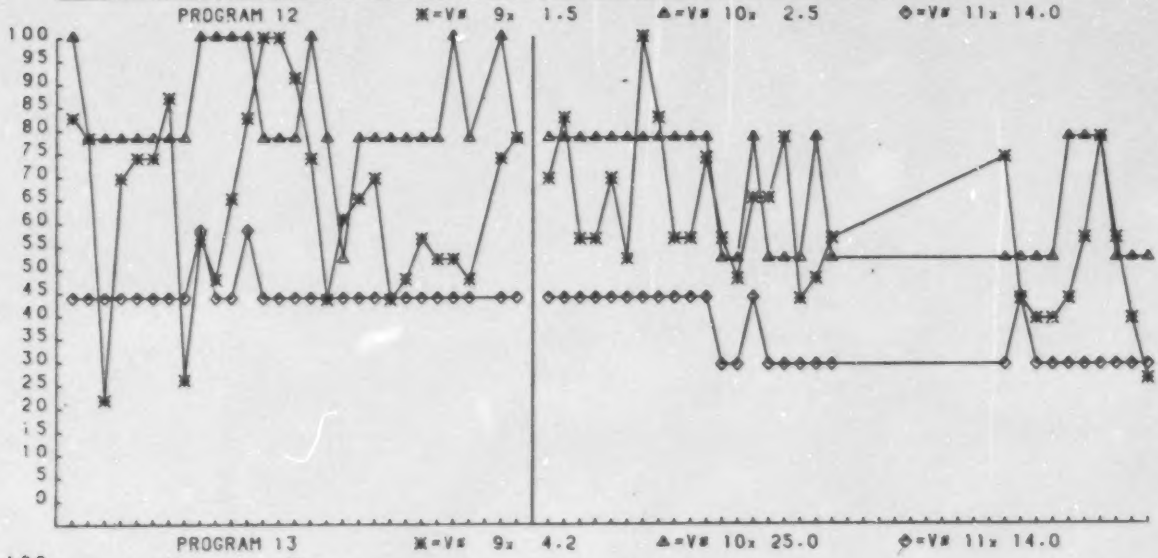

PROGRAM 15
*=V# 9x 10.0
▲=V# 10x .1
◇=V# 11x 14.0

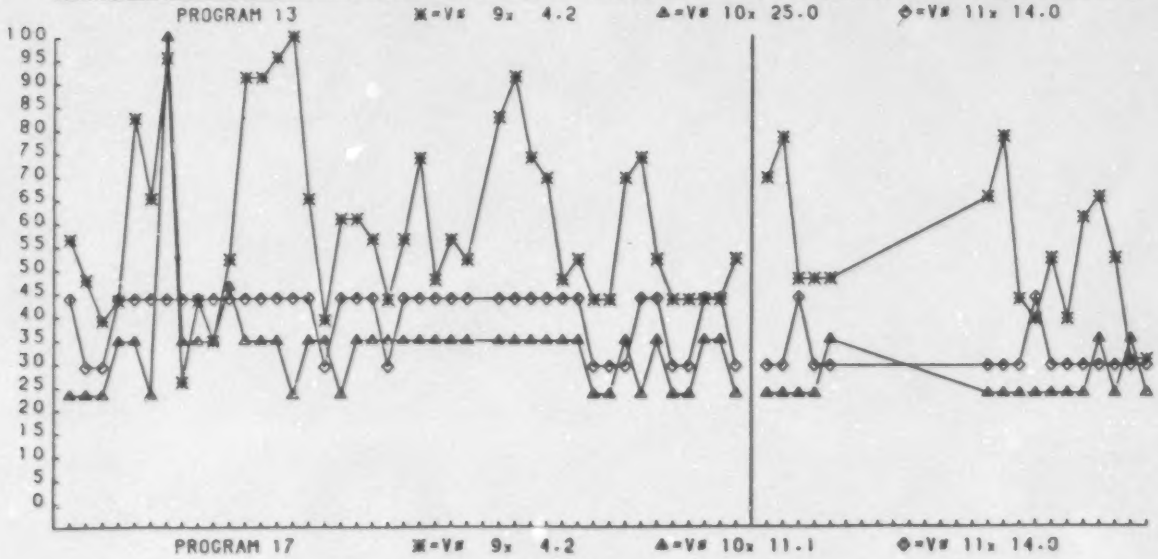

PROGRAM 13
*=V# 9x 10.0
▲=V# 10x .3
◇=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0
▲=V# 12x 20.0
◇=V# 13x 16.7

FAMILY 157

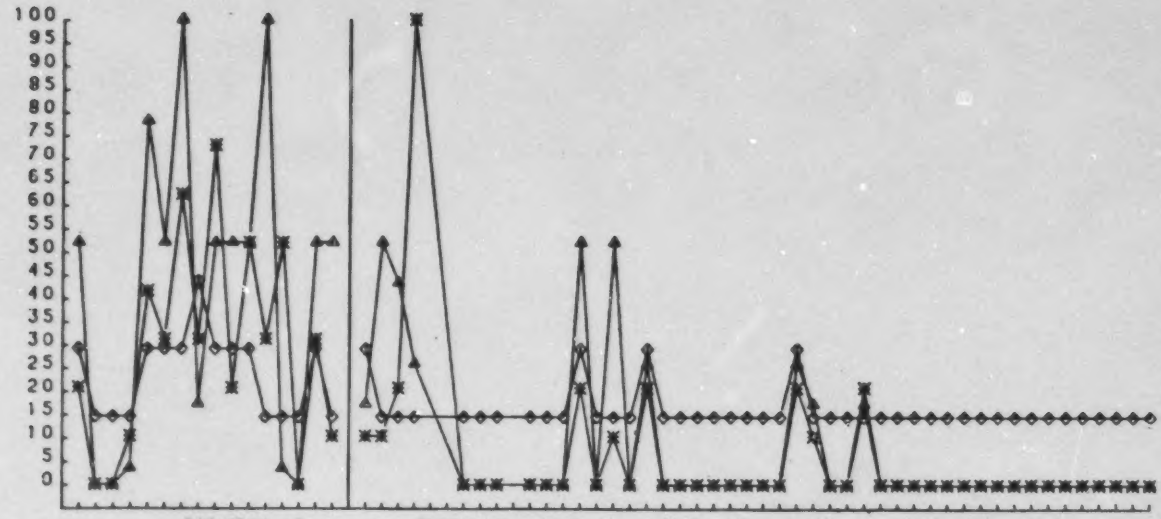


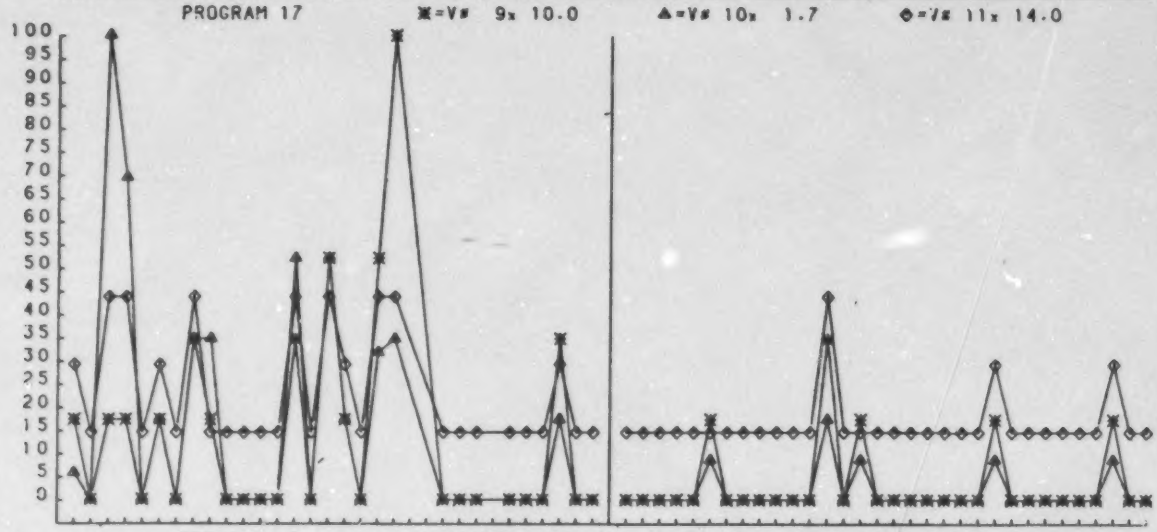


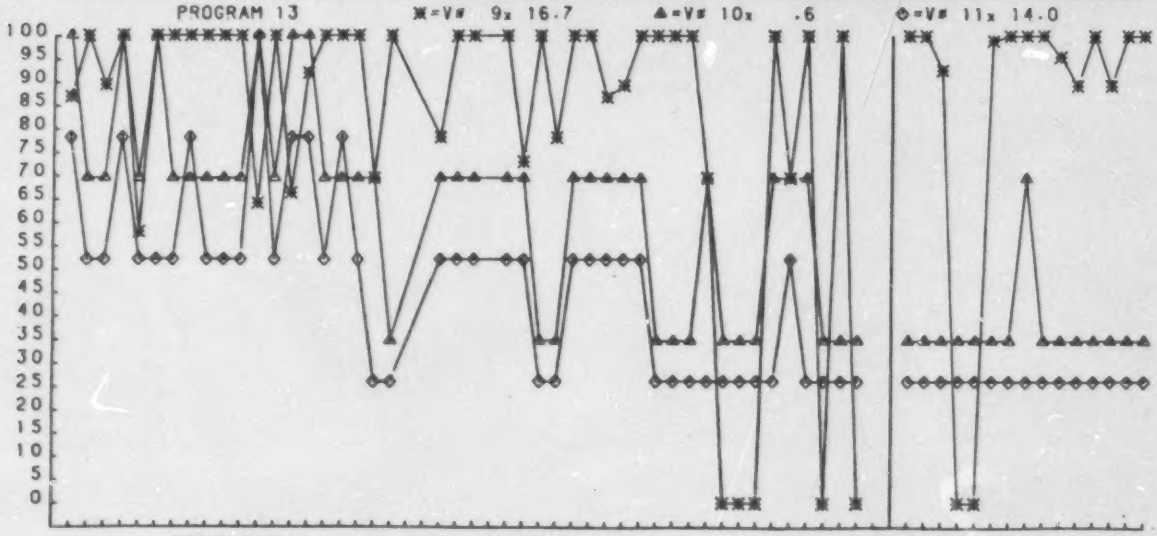


FAMILY 158

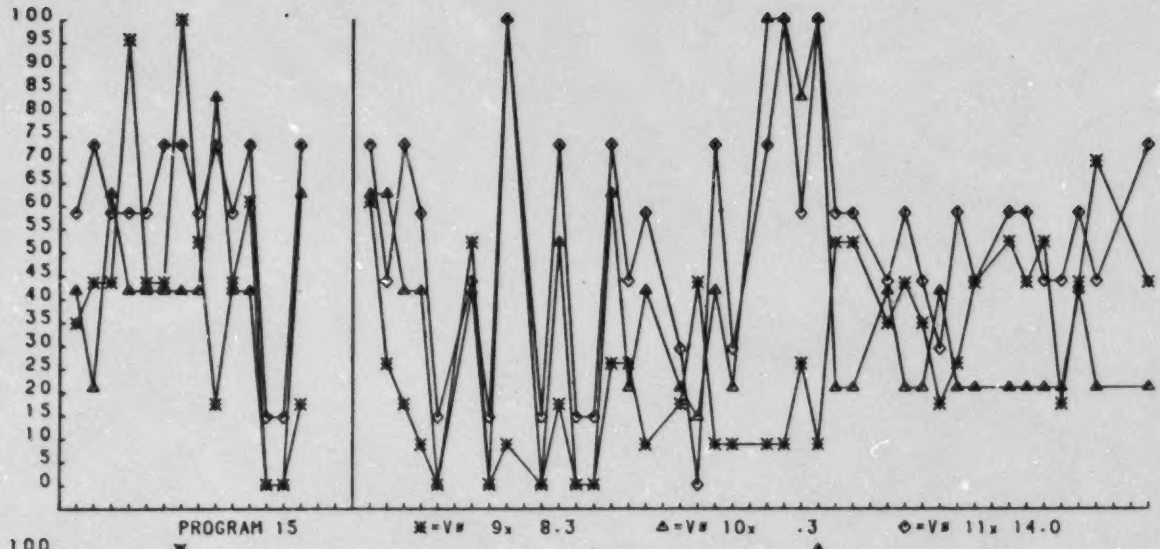
R74

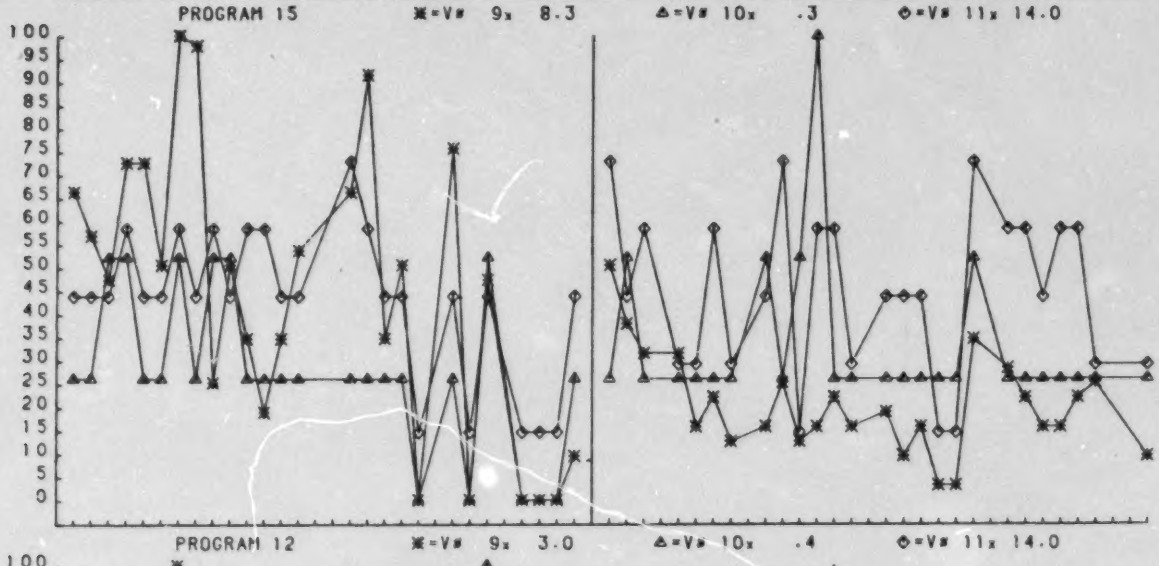

PROGRAM 17
*=V# 9x 10.0
▲=V# 10x 1.7
◆=V# 11x 14.0

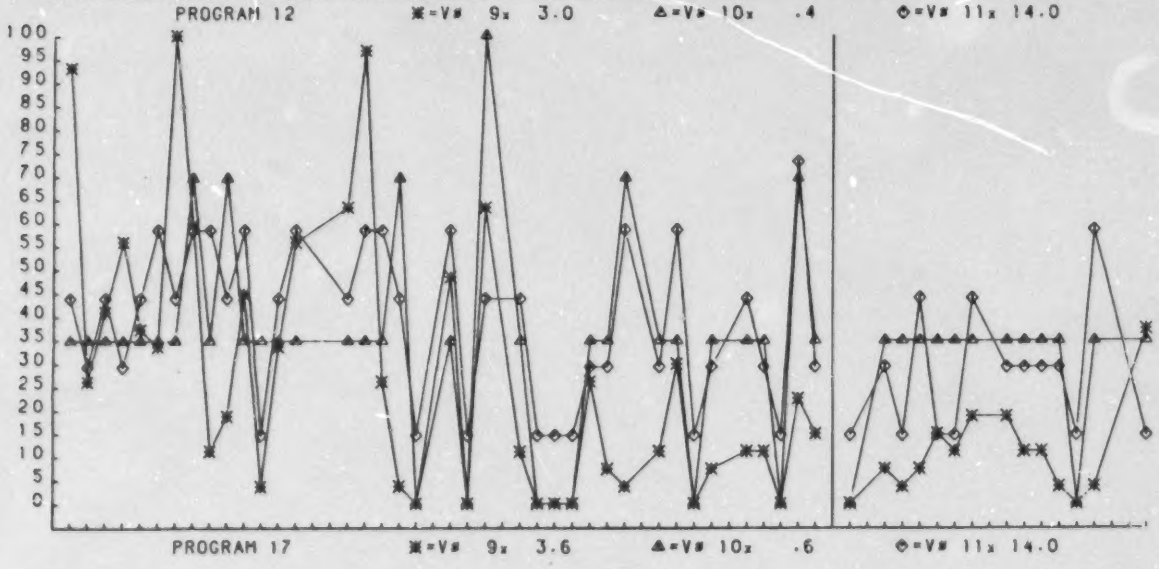

PROGRAM 13
*=V# 9x 16.7
▲=V# 10x .6
◆=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0
▲=V# 12x 33.3
◆=V# 13x 25.0

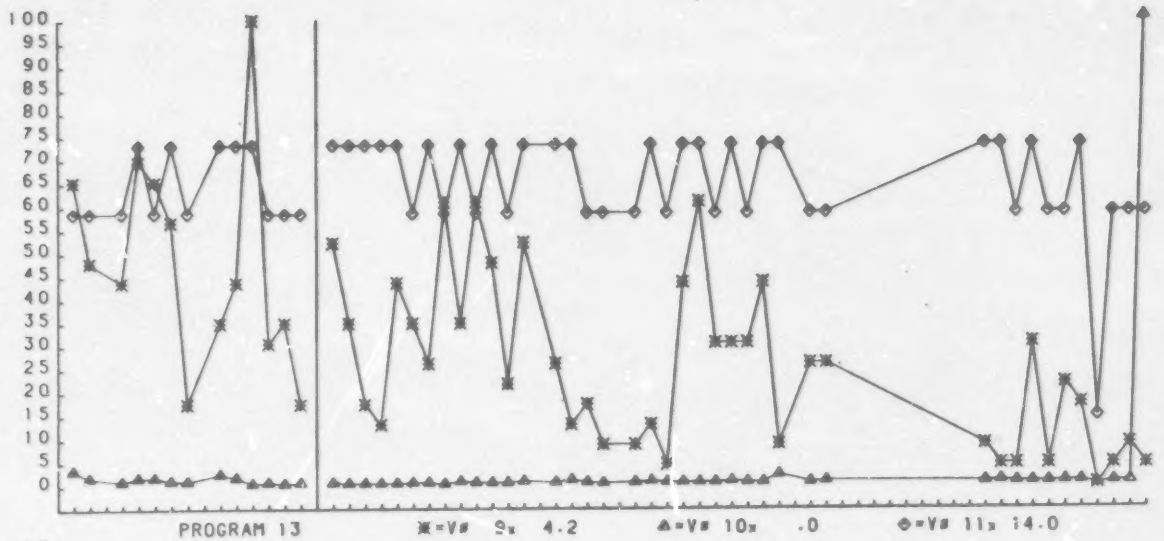
FAMILY 159

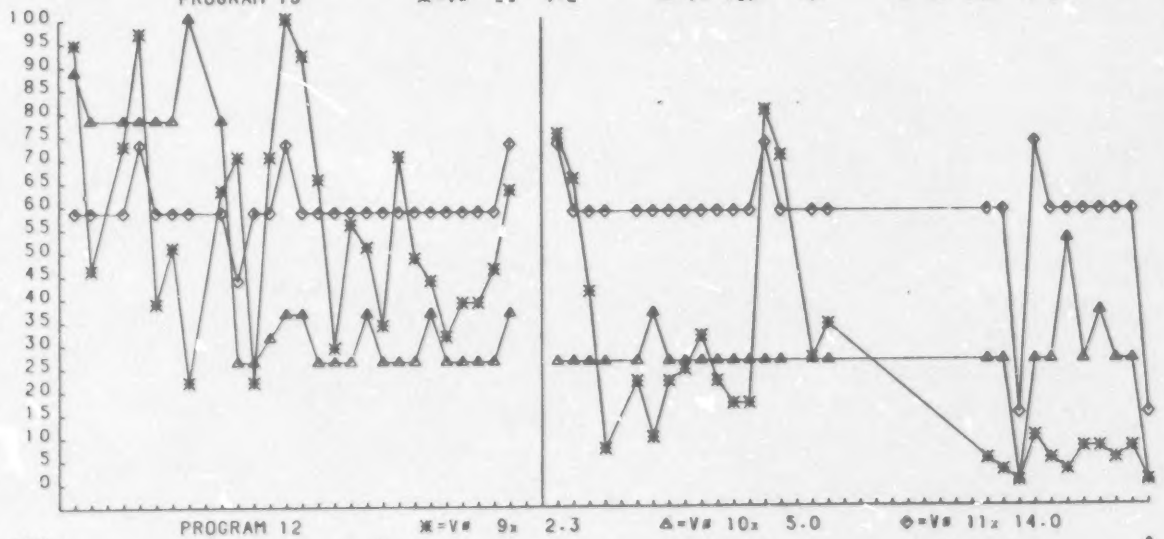

PROGRAM 15
*=V# 9x 8.3
△=V# 10x .3
◇=V# 11x 14.0

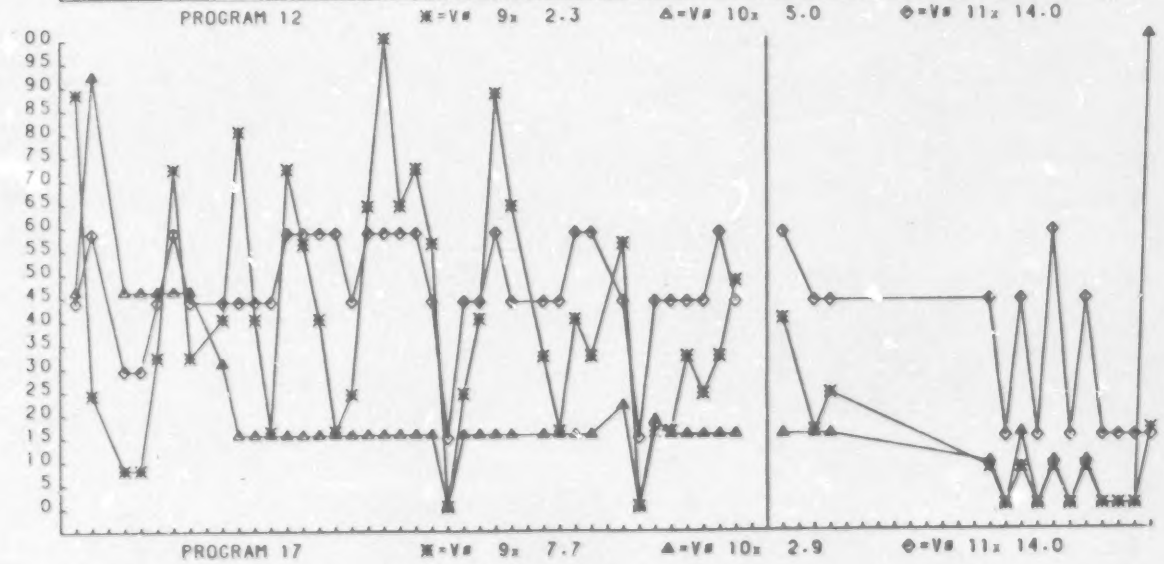

PROGRAM 12
*=V# 9x 3.0
△=V# 10x .4
◇=V# 11x 14.0


PROGRAM 17
*=V# 9x 3.6
△=V# 10x .6
◇=V# 11x 14.0

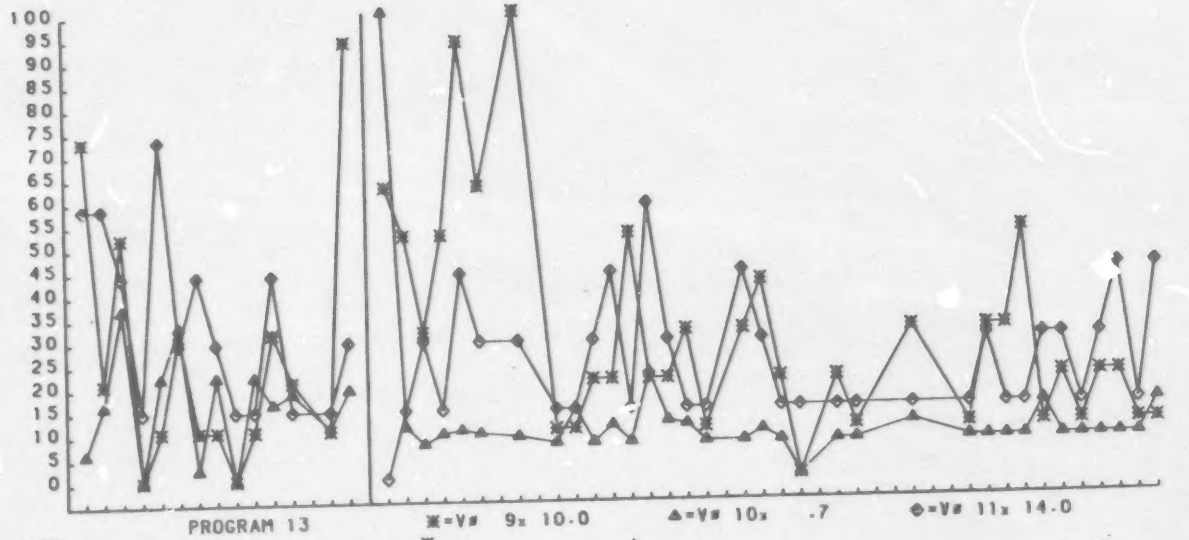
FAMILY 160

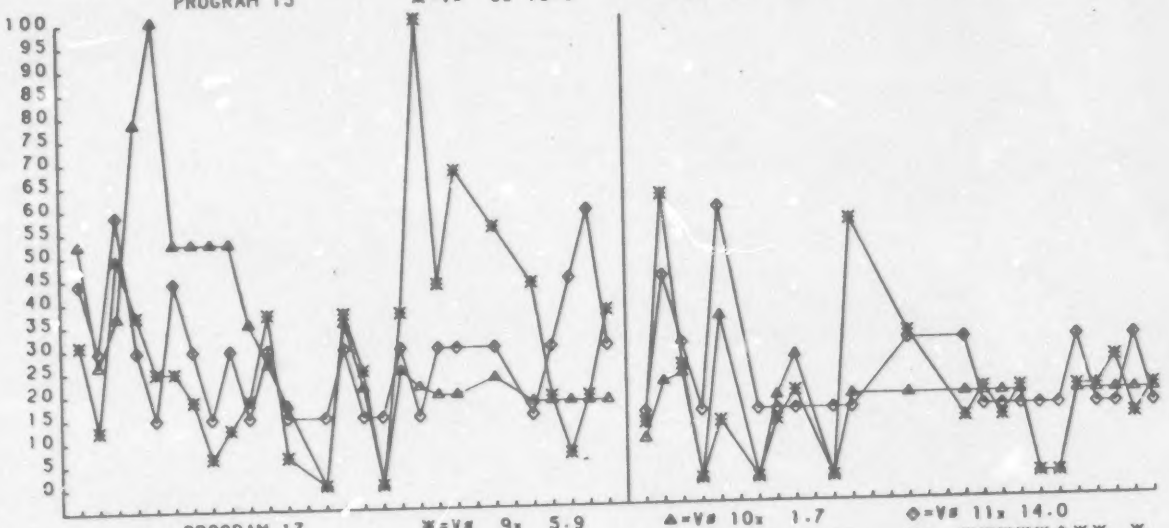

PROGRAM 13
*=V# 9x 4.2
▲=V# 10x .0
◇=V# 11x 14.0

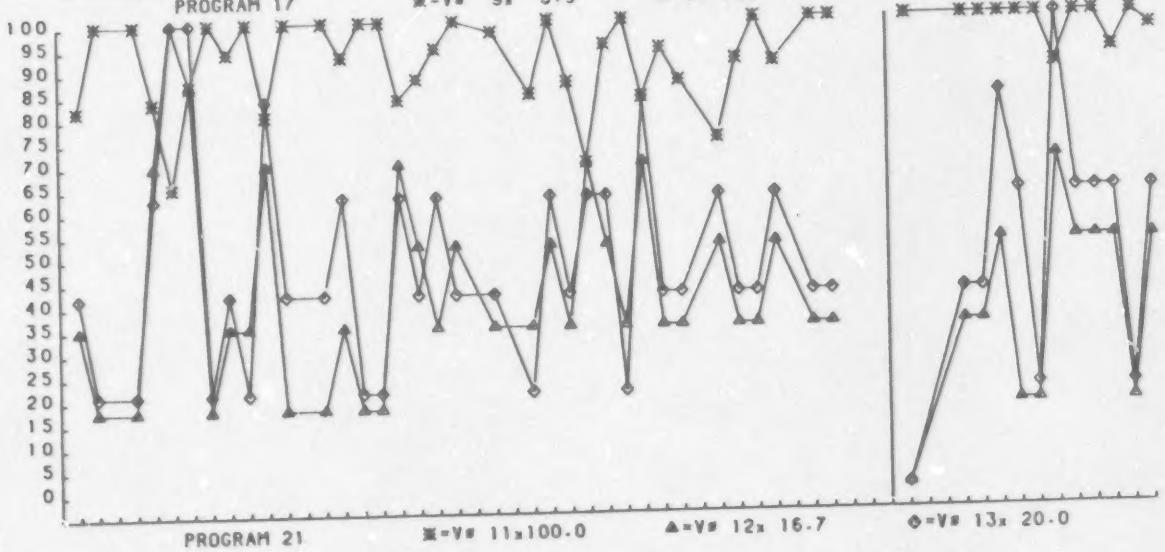

PROGRAM 12
*=V# 9x 2.3
▲=V# 10x 5.0
◇=V# 11x 14.0


PROGRAM 17
*=V# 9x 7.7
▲=V# 10x 2.9
◇=V# 11x 14.0

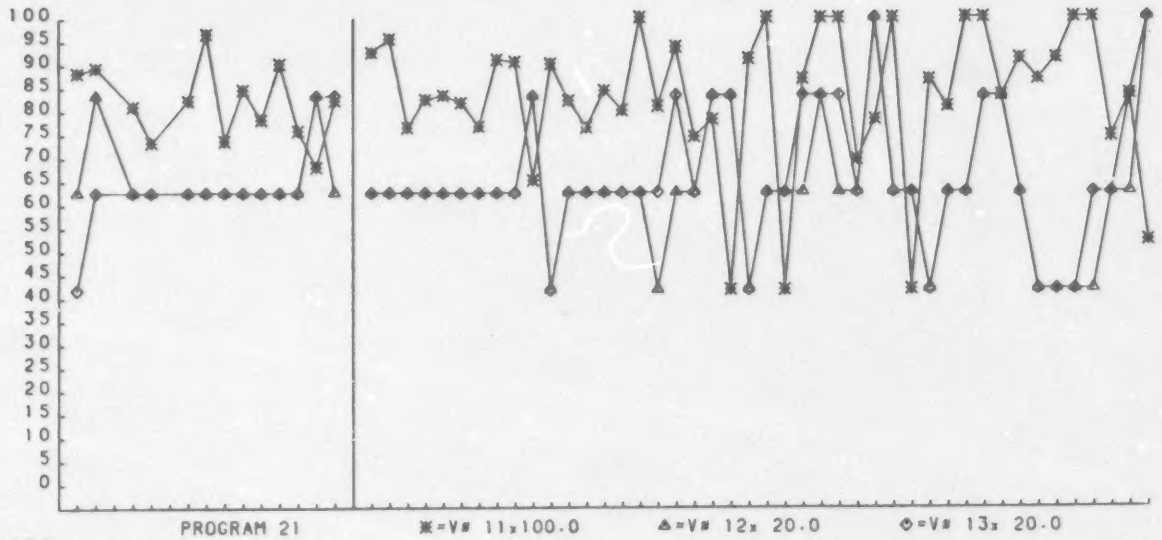
FAMILY 162

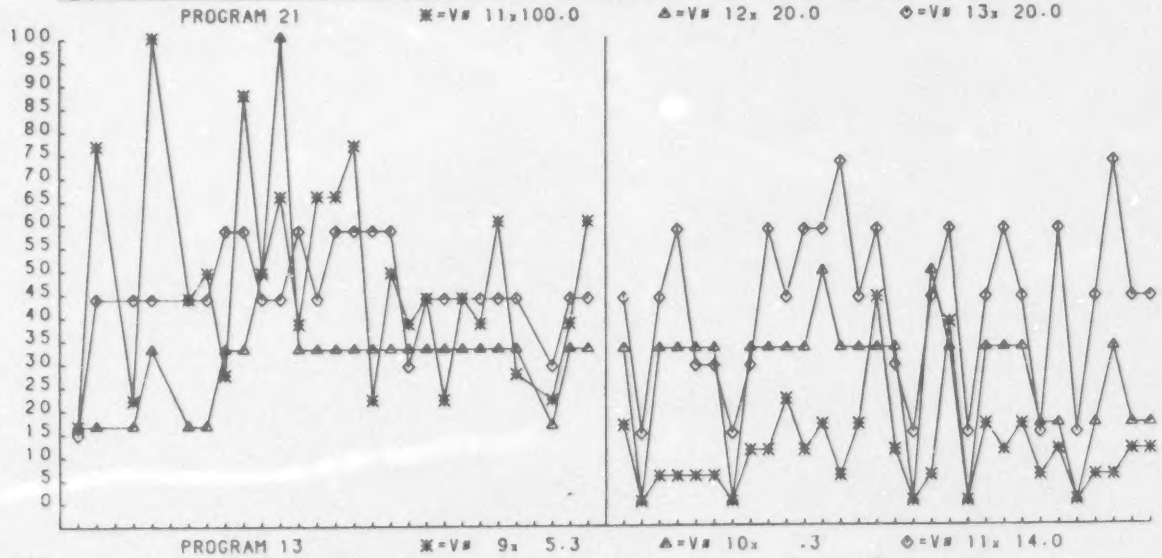

PROGRAM 13
*=V# 9x 10.0
▲=V# 10x .7
◆=V# 11x 14.0

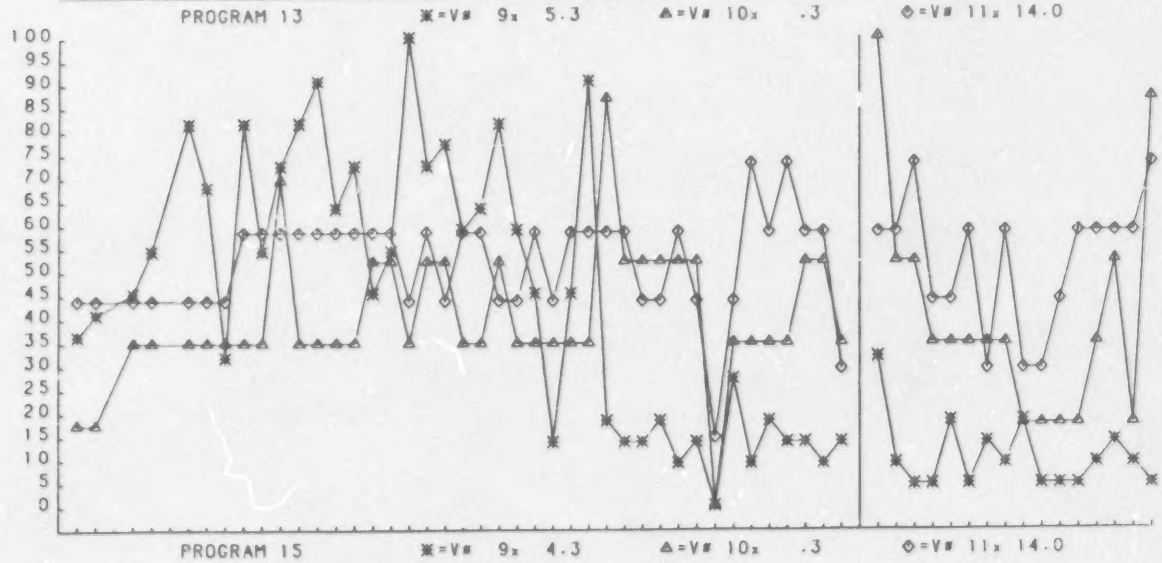

PROGRAM 17
*=V# 9x 5.9
▲=V# 10x 1.7
◆=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0
▲=V# 12x 16.7
◆=V# 13x 20.0

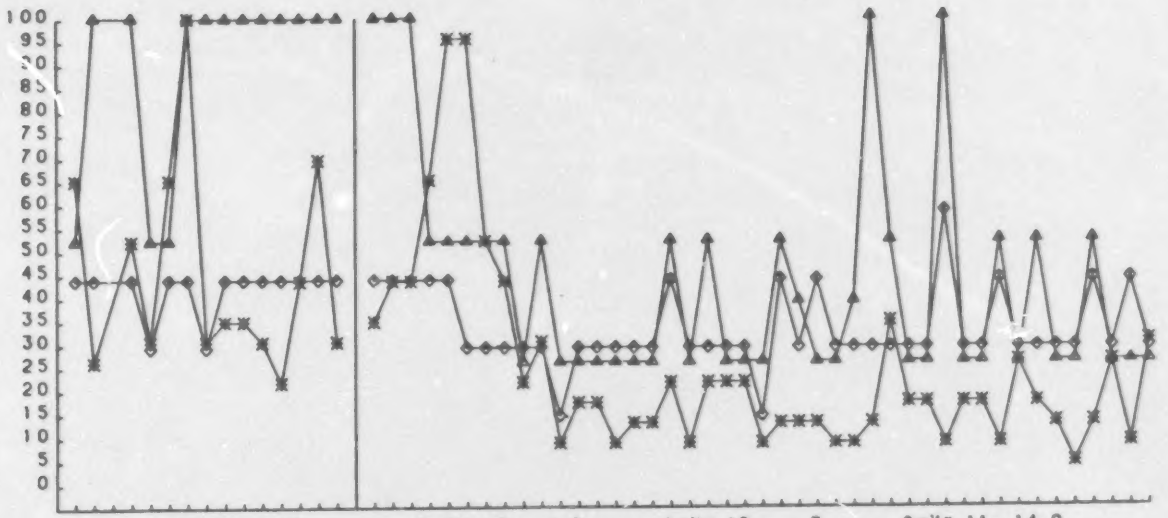
FAMILY 163

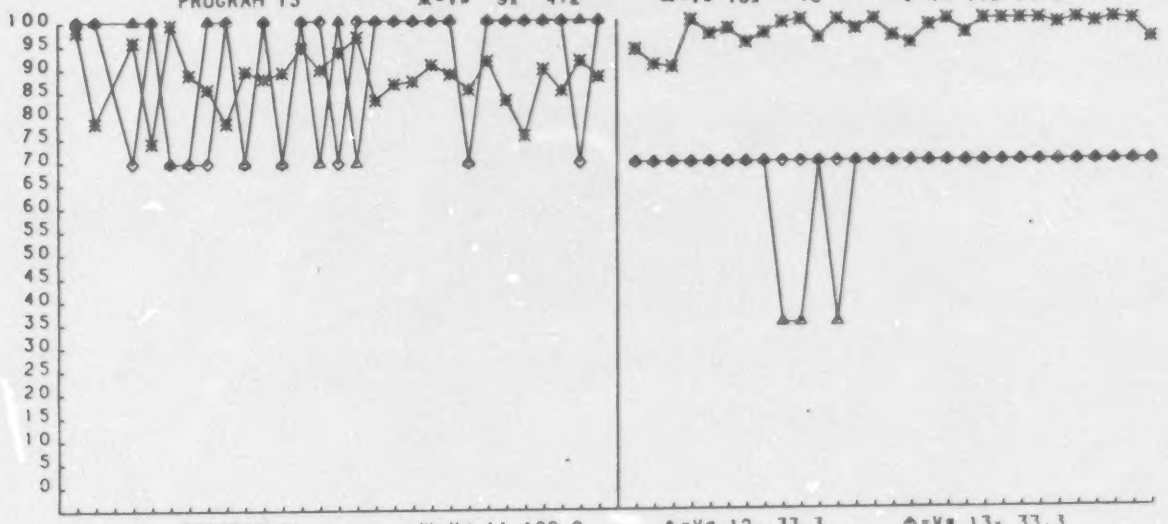

PROGRAM 21
*=V# 11x100.0
△=V# 12x 20.0
◇=V# 13x 20.0

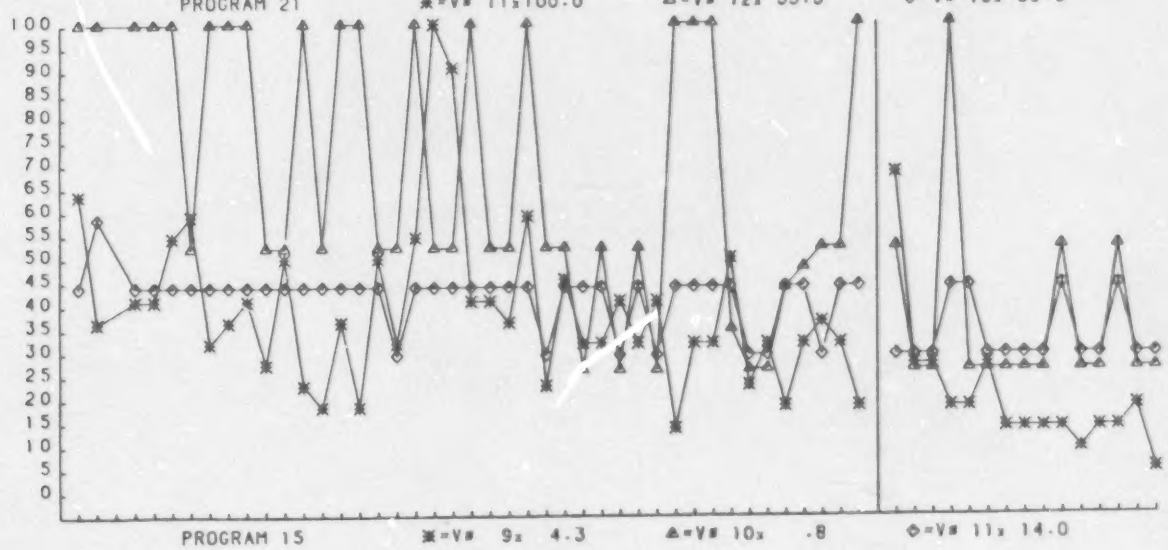

PROGRAM 13
*=V# 9x 5.3
△=V# 10x .3
◇=V# 11x 14.0


PROGRAM 15
*=V# 9x 4.3
△=V# 10x .3
◇=V# 11x 14.0

FAMILY 168


PROGRAM 13
*=V# 9x 4.2
△=V# 10x .8
◇=V# 11x 14.0


PROGRAM 21
*=V# 11x100.0
△=V# 12x 33.3
◇=V# 13x 33.3


PROGRAM 15
*=V# 9x 4.3
△=V# 10x .8
◇=V# 11x 14.0

FAMILY 169

Approximately 47 of the 72 evaluations, as judged from these figures, appear to demonstrate internal validity, that is, evidence that changes produced were produced by the programs implemented rather than by artifacts. Our estimates of internal validity are based on changes in slope or mean occurring at or soon after a treatment was implemented rather than at other times. Graphs which show no change, which indicate an undesirable trend, or which have excessive variability are also assumed to lack support for internal validity; approximately 25 evaluations had one or more of these "faults" leading to our estimate of 47/72 successful evaluations.

In each figure, the major measure of outcome, and the one on which other analyses are based, is indicated with an asterisk. In addition, families were sometimes derelict in collecting or reporting data, so some graphs have days with incomplete data.

To allow actuarial summation across subjects and programs (for descriptive purposes only) a measure called "Percentage of possible change" was computed for the major behavior change measure for each program. This measure is computed by estimating the size of maximum possible change in the clinically desirable direction. For instance, if the baseline mean score for following directions for a particular child was 55% directions obeyed, he could improve 45% before reaching the theoretical limit of 100% compliance. If, in fact, his treatment mean was 85%, indicating a 30% improvement, his percent of possible change as 30/45 x 100 or 67%. Similarly, if a child had a baseline mean for interrupting of 25 interruptions, the maximum improvement possible would be a decline to zero. If, in fact interrupting decreased to 10 (treatment mean), his percent of possible maximum change would be 60%, that is, 15/25 x 100.

THIRD YEAR SELF-REPORT SUBJECTS

Percentage Change of Maximum Possible Change

(Median Change = 45%)

	-95 to -81	-94 to -80	-79 to -65	-64 to -50	-49 to -35	-34 to -20	-19 to -5	-4 to 10	11 to 25	26 to 40	41 to 55	56 to 70	71 to 85	86 to 100	Total
Following Directions							2	1	3	2	3	3	2	1	17
Whining	1								1	1	2	3	1		9
Fighting							3	1		2	4	2	4		16
Tantrums		1			1					2	4	1	2		11
Interrupting	1								3	2	2	2	2		12
Bedtime								1	1		2				4
Room Neatness								1	1				1		3
TOTAL	2	1	0	0	1	0	5	4	9	9	17	11	12	1	72

These scores are merely a convenience. In many respects they are also conservative, as no chilren are perfect, and few individuals would argue that no "undesirable" behavior is a reasonable clinical goal. In addition, this method of computation provides a limit of 100% for desirable change (ceiling effect), but no limit for undesirable change. Thus, a child with baseline fighting of 3 instances per day who increased in treatment to 9 per day would obtain a score of -300%. Finally, it is computed on the entire treatment period, which varied in length depending when a program was implemented. However, some treatments may not produce immediate results, and effects may only appear after time.

The frequency disistribution of changes as measured by this score is presented in the table which follows for individual programs and for the total population of evaluations across subjects and programs. If the category closest to zero (-4% to +10%) is taken as "no change" and all other categories are assumed to indicate some positive or negative change, even if only descriptively, we note that 4 of 72 evaluations indicate no change, 9 of 72 indicate a negative change, and 59 of 72 indicate a positive change, with a median change of +46 %. By and large these oucomes, using self-instructional materials which would cost about $1.00 in clinical use, indicate changes comparable to those reported in assessments of outcomes of child treatment by usual clinical procedures.

It should also be noted that the number of evaluations for each of the seven programs varied significantly, as parents selected freely the programs they wished to use. Thus, there are 17 separate evaluations of the following directions, program and 16 of fighting, while for bedtime and room neatness there are only 4 and 3 respectively. The variations of percent change within each program do not indicate any glaring

differences in differential effectiveness, although for those two programs with few evaluations conclusions are difficult to draw.

Combined Reliability and Validity

Parent and observer data provided both reliability and validity measures, with the latter interpreted as measures of concurrent validity. Two types of data were collected: (1) those based on parent recall of the number of behaviors following the hour of observation (collected for all families) and (2) those based on an interval by interval basis during which the parent signalled the occurrence of each instance of a given behavior (collected on 12 families). Here the observer recorded not only the parent's indications but also what he/she observed independently of the parent.

In addition, during this year and prior years interobserver reliability on weighted percentage occurrence and non-occurrence was taken with separate assessments for each score during baseline and treatment. These reliabilities were consistently above +.70 (except during early R & D periods during year one). The analyses which follow are not based upon these reliabilities, but on the relationships between observer scores and parent records.

Reliability/Validity Based on Recall

This is presented separately for two groups because of procedural differences which may have affected these outcomes (See table on next page). Groups A and B contained 12 families each. Group B also had interval data dollected. Pearson's Correlation Coefficcient were compiled between parent recall data and observer record data. The degrees of freedom ranged from 4 to 16.

The nine coefficients for the Following Directions Program for Group A ranged between .08 and .88, with a median of .37. The eight

RECALL DATA

Summary of Correlational Analyses for Groups by Programs, using Range and Median

Groups	Programs	Range	Median	No. of significant correlations ($p \leq .05$)
A	Following Directions	.08 to .88	.37	2
B	Following Directions	.16 to .85	.40	2
A	Tantrums	.61 to .88	.69	2
B	Tantrums	.27 to .97	.48	4
A	Fighting	.19 to .89	.52	2
B	Fighting	.24 to .97	.64	2
A	Whining	.18 to .70	.62	2
B	Whining	.11 to .61	.51	0
A	Interrupting	.07 to .65	.60	0
B	Interrupting	.07 to .71	.46	1

coefficients for Group B range between .16 and .85, with a median of .40 (positive and negative signs were ignored in computing all the medians). Group A had two significant correlations, Group B two.

The four coefficients for the Tantrums program for Group A ranged between .61 and .88, with a median of .69. The seven coefficients for group B ranged between .27 and .97, with a median of .48. Group A had two significant correlations, Group B four.

The seven coefficients for the Fighting Program for Group A ranged between .19 and .89, with a median of .52. The nine coefficients for Group B ranged between .24 and .97, with a median of .64. Group A had two significant correlations, Group B two.

The four coefficients for the Whining Program for Group A ranged between .18 and .70, with a median of .62. The five coeffcients for Group B ranged between .11 and .61, with a median of .51. Group A had two significant correlations, Group B none.

The five coefficients for the Interrupting Program for Group A ranged between .07 and . .65, with a median of .6. The six coefficients for Group B ranged between .07 and .71, with a median of .46. Group A had no significant correlation, Group B had one.

Reliability and Validity Based on Interval Data

The Lahey-Harris formula was used to compute percent agreement on the within session interval data. This applied to Group B only. The index was based on parent's indication of occurrence or non-occurrence of behavior and observer's interval recording. Two sets of agreement coefficients were computed because of the highly stringent and conservative nature of the Lahey-Harris when analyzing low frequencies. Thus, agreement coefficients were computed separately for sessions in which both parent and observer indicated frequencies greater than three

and for total sessions regardless of observed or reported frequencies.

In the Following Directions program directions to "stop" a behavior were recorded separately from directions to "start" a behavior, and whether or not a child obeyed each of these types of directions was also recorded separately. Thus, percent agreement coefficients for 4 different measures are provided. Seven families provided 15 data sets on which to compute percent agreement. When low frequencies were exclued the percent agreements were 90% for start directions given, 84% for start directions obeyed, 100% for stop directions given, and 100% for stop directions obeyed. When low frequencies were not excluded the percent agreements were 86%, 72%, 83% and 69% for the same respective measures.

For the tantrums program, percent agreement of frequency of tantrums was 87% if the low frequency sessions are omitted, and 62% when these sessions are included. For fighting, the respective percent agreements are 71% and 67%. For whining, which had the lowest percent agreements, the agreement was 59% with low frequency sessions omitted and 55% when these were included. For interrupting the percent agreement was 91% with low frequency sessions omitted and 64% when these were not excluded. The table of average percent agreements summarizes this information as well as indicating the number of families and assessments on which each percent agreement is based.

INTERVAL DATA

Summary of Average Percent Agreements and Related Data

Program	Score	Number of Families	Number of Assessments	Without Low Frequencies	With Low Frequencies
Following Directions	start given	7	15	90%	86%
	start obeyed	7	15	84%	72%
	stop given	7	15	100%	83%
	stop obeyed	7	15	100%	69%
Tantrums	frequency	6	18	87%	62%
Fighting	frequency	8	25	71%	67%
Whining	frequency	3	7	59%	55%
Interrupting	frequency	6	16	91%	64%

CHAPTER SIX

Year Three Mail-In Subjects

Nine families were accepted for a mail-in only procedure in year three. These families simply mailed in their reports on a weekly basis, using the same self-report measures as the year three self-report families. Families were placed into this group primarily because of the distance factor (over 15 miles away from the University). However, we were interested in whether parent data using the revised self-report scale with no staff contact would show effects related to treatment onset, or general effects (as in year two) unrelated to treatment onset. It was felt that correlation between treatment onset and changes in the self-report measures would be an indirect indication of validity. No reliabilities of any sort were obtained.

Intake Characteristics

Eight of the 9 families in the mail-in group were two-parent families, in one the parents were divorced. All fathers were employed, five in professional, technical, administrative or managerial positions, two in sales or clerical work, one in a laboring job and one in a service position. Eight of the others were housewives not otherwise employed, while one mother was a sales or clerical worker. Of the seven families who replied to the question on family income, a mean family income of $21,486 was reported, with a range of $13,400 to $25,000.

The average father was 32 years old and had completed 14.5 years of school, while the average mother was 31 years old with 13.8 years of schooling completed.

Children ranged in age from 44 to 92 months, with an average of 65

months. Six of the children were male, three female. The median number of siblings was two.

This group was selected for the less structured mail-in procedure because they did not meet our usual criteria concerning handicap or severity of behavior problem, or because they lived too far away to be serviced in any other group. Seven of the children were listed by parents as having no apparent handicap, two were listed by parents as learning disabled or having multiple handicaps. Parents reported that these statuses had been verified by a physician. Seven of these nine chilren currently attended some outside setting during the day, and six had attended at least one previous setting. Most reported that this setting was not a special program.

Cognitive Tests

On cognitive tests, which assessed the parent's answers to written multiple choice questions based upon the programs, the maximum possible score attainable on each test was ten. For all tests the parents in this group attained a mean score of 8.4, with typically perfect or near perfect scores for all but one parent. This suggests that for all but one of the families "verbal" comprehension of the programs was satisfactory, and problems did not arise out of inability to read and comprehend the programs.

Results

A summary of results in terms of percent of possible change is presented in the table labeled "Third Year Mail-In Subjects." Seven families completed data collection, providing 21 evaluations of individual programs. Fighting was the most commonly used program with this group, with interrupting, following directions and whining selected

THIRD YEAR MAIL-IN SUBJECTS

Percentage Change of Maximum Possible Change

	-34 to -20	-19 to -5	-4 to 10	11 to 25	26 to 40	41 to 55	56 to 70	71 to 85	86 to 100	TOTAL
Following Directons			3	1						4
Whining					1		1	1		3
Fighting				2	3	2				7
Tantrums				1		1				2
Interrupting					4					4
Room Neatness	1									1
TOTALS	1	0	3	4	8	3	1	1	0	21

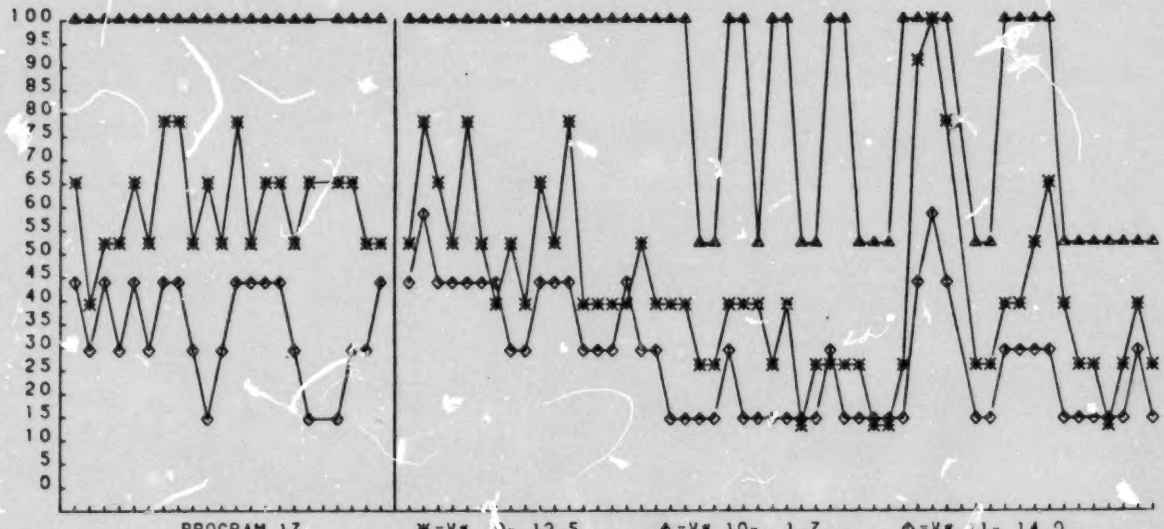
Median = 31%

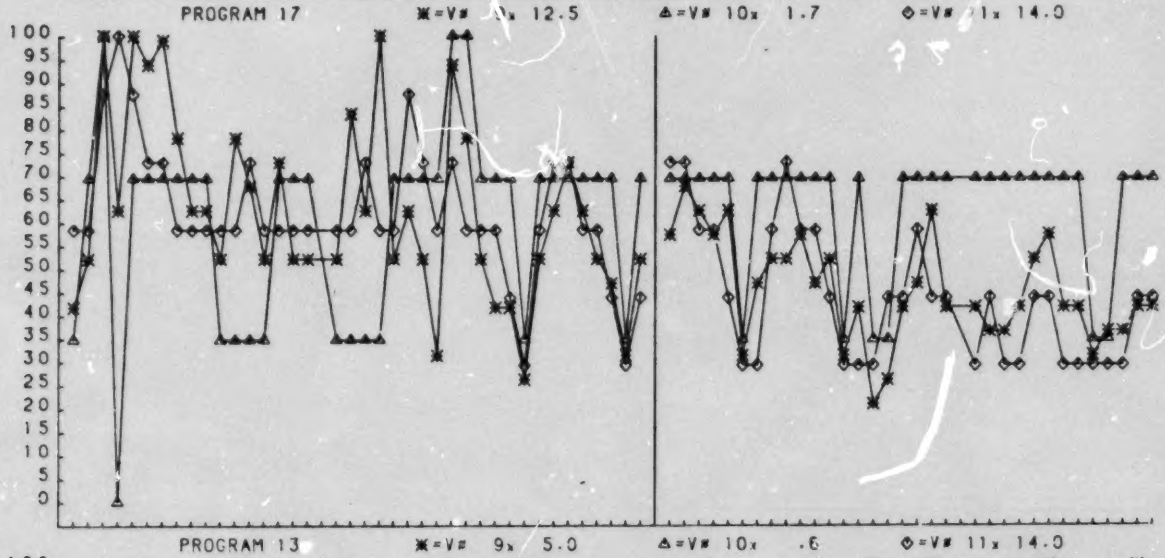
next most frequently. Problems related to behavioral excesses (fighting, whining and interrupting) rather than deficits (following directions) were more common in this group while the reverse was true of the main year three group. This may reflect the selection factors that lead to placement in the mail-in group, particularly less serious indication of handicap and apparently less severe behavior problems.

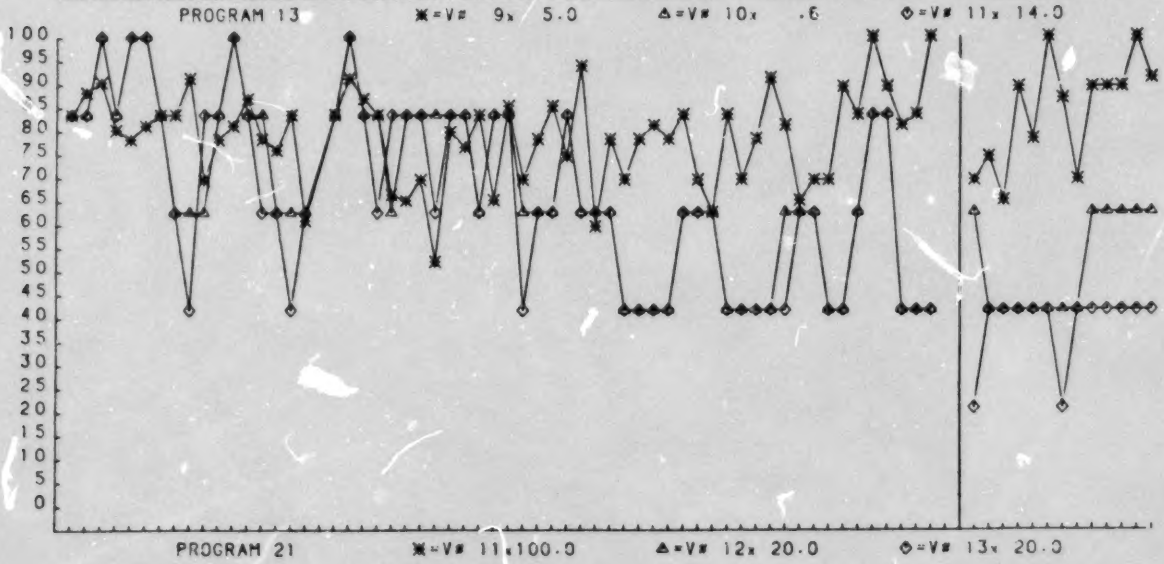
As the table cited indicates, 17 of the 21 evaluations resulted in positive change, and one indicates a negative change, with a median change of 31% of possible change.

Figures 6-1 to 6-7 present the individual graphs. Again, attention should paid to the variables graphed with asterisks; these are the major outcomes variables, and indicate either frequency of a behavior or percent compliance, tasks completed, etc. In family 132 effects start with treatment onset. In family 135 and in family 136 this is true, except a clear effect is not visible for one program in each family. In family 151 note that the baseline level for following directions was so high that a ceiling effect exists, and it is difficult to demonstrate any treatment effect. In families 152 and 155 one treatment in each shows effects which start before treatment onset.

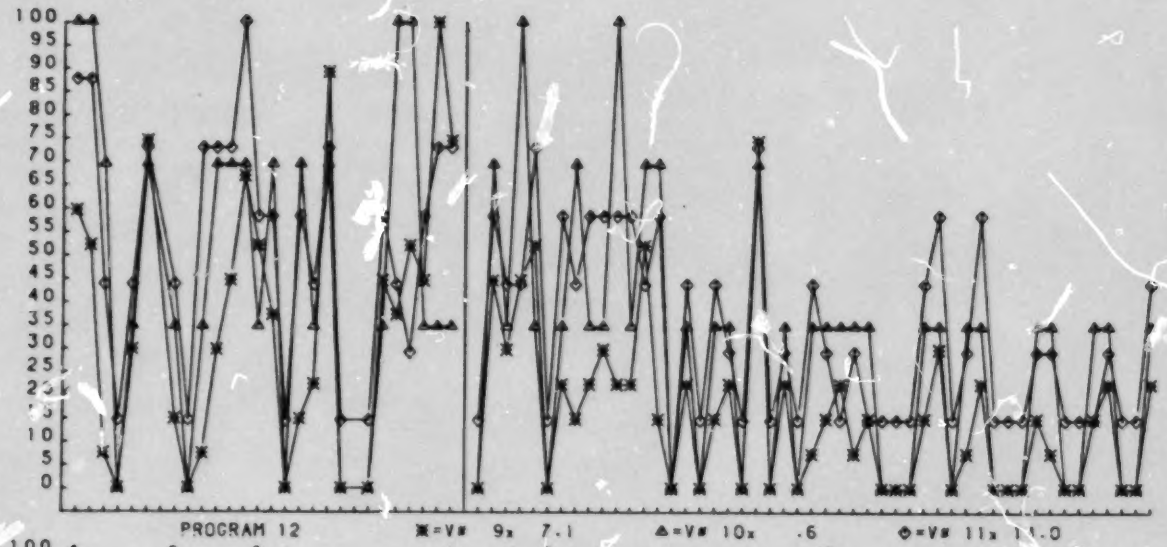
Although generalized conclusions are impossible, it is interesting to note that this group appears to have had less success with the popular following directions program than the main year three group, and less success in general. As figures 6-1 to 6-7 indicate, baseline rates of compliance in this program were high, probably already in the "normal" range for all but one family. This ceiling effect may account for the lower success in terms of magnitude of percent of possible change.

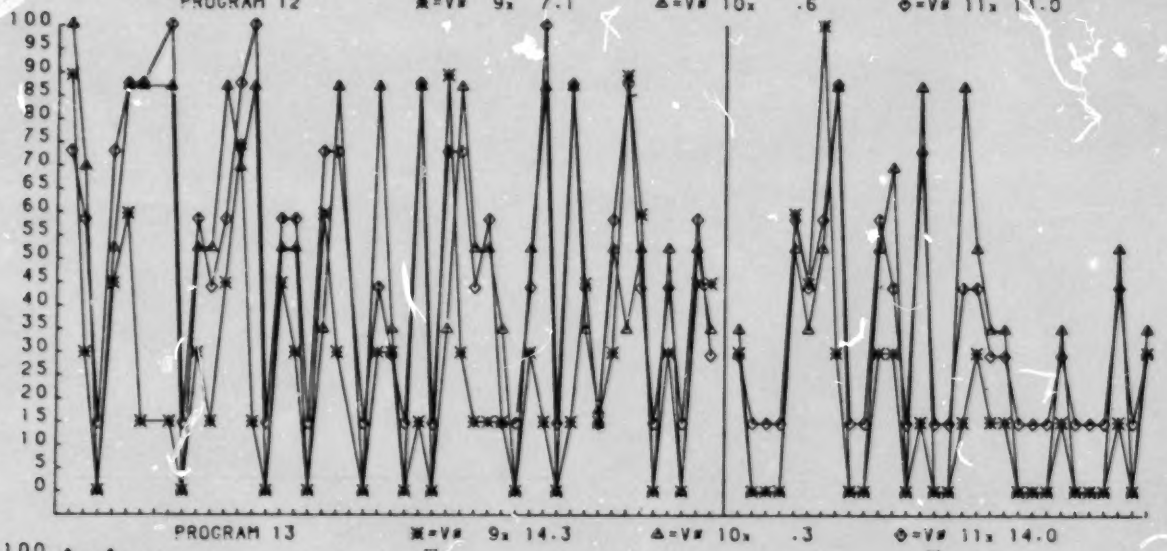

PROGRAM 17
✱=V# 9x 12.5
△=V# 10x 1.7
◇=V# 11x 14.0

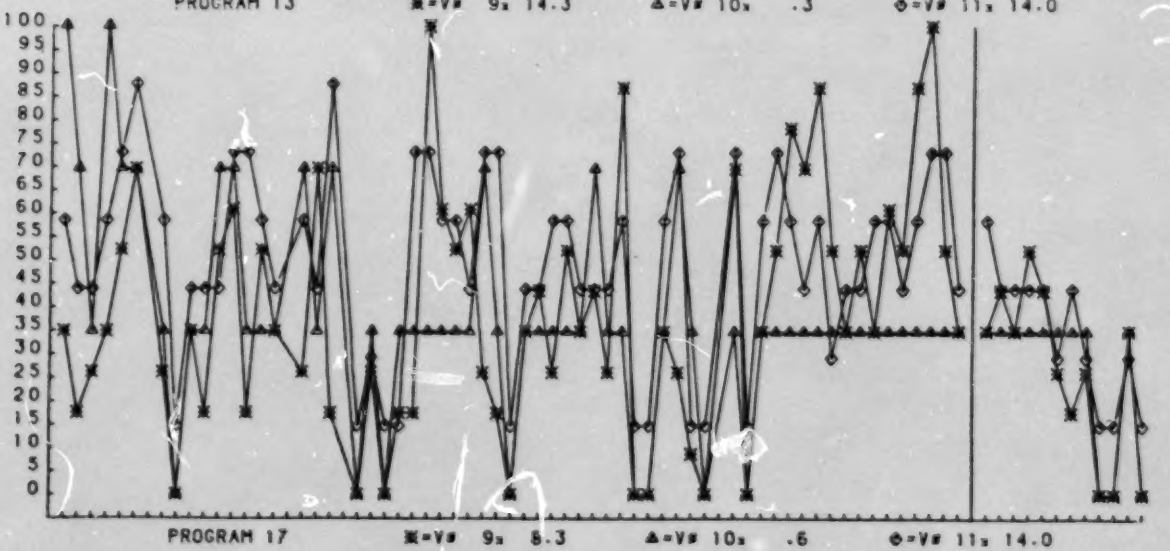

PROGRAM 13
✱=V# 9x 5.0
△=V# 10x .6
◇=V# 11x 14.0


PROGRAM 21
✱=V# 11x100.0
△=V# 12x 20.0
◇=V# 13x 20.0

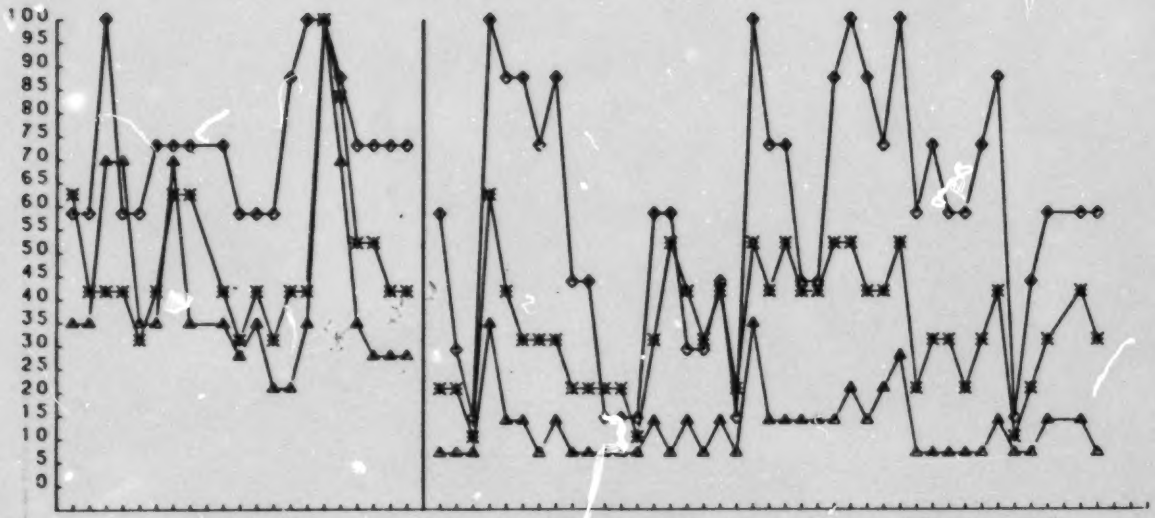
FAMILY 132

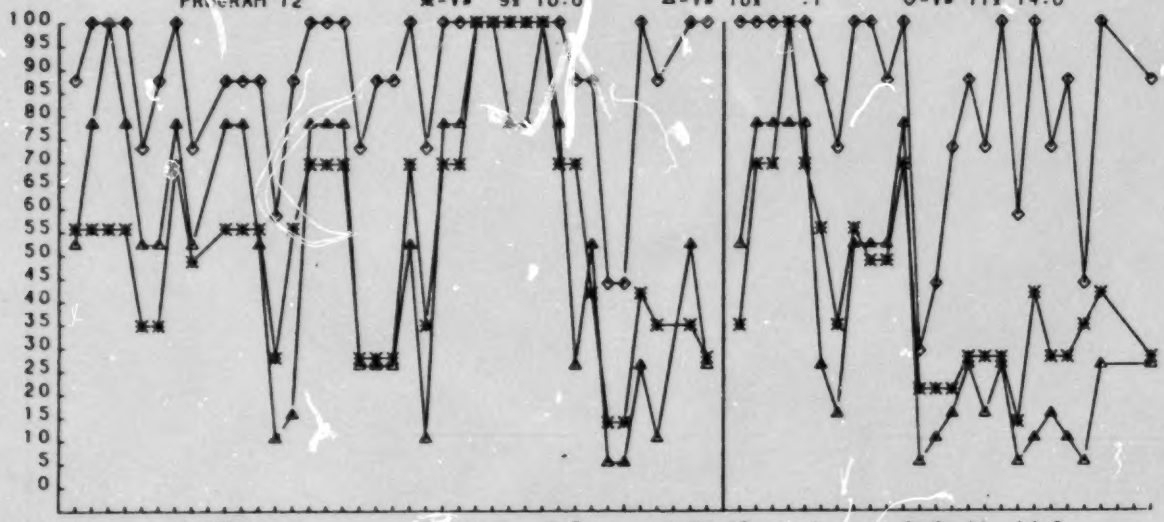

PROGRAM 12
*=V# 9x 7.1
Δ=V# 10x .6
◇=V# 11x 14.0

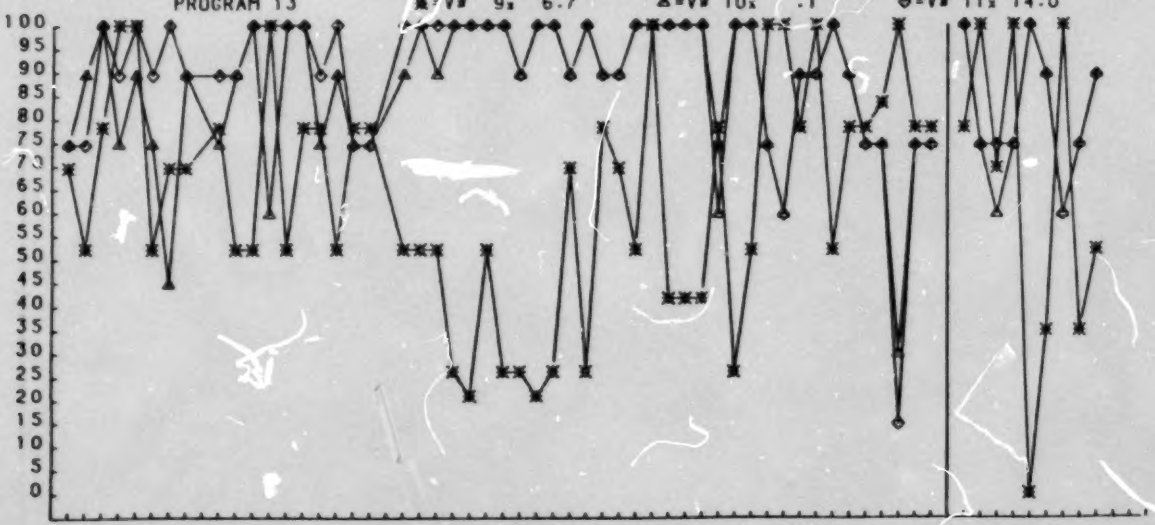

PROGRAM 13
*=V# 9x 14.3
Δ=V# 10x .3
◇=V# 11x 14.0


PROGRAM 17
*=V# 9x 8.3
Δ=V# 10x .6
◇=V# 11x 14.0

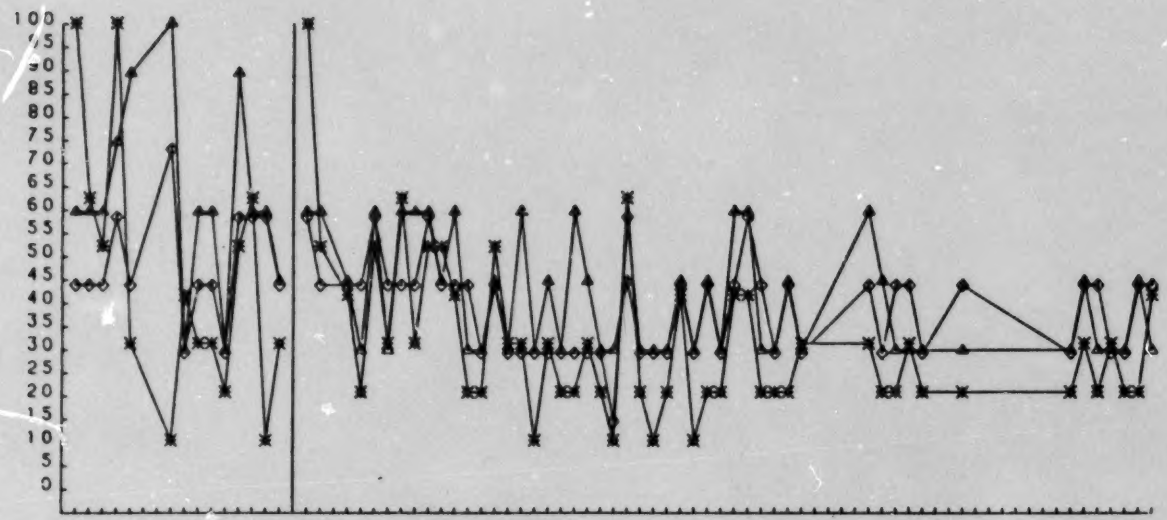
FAMILY 135

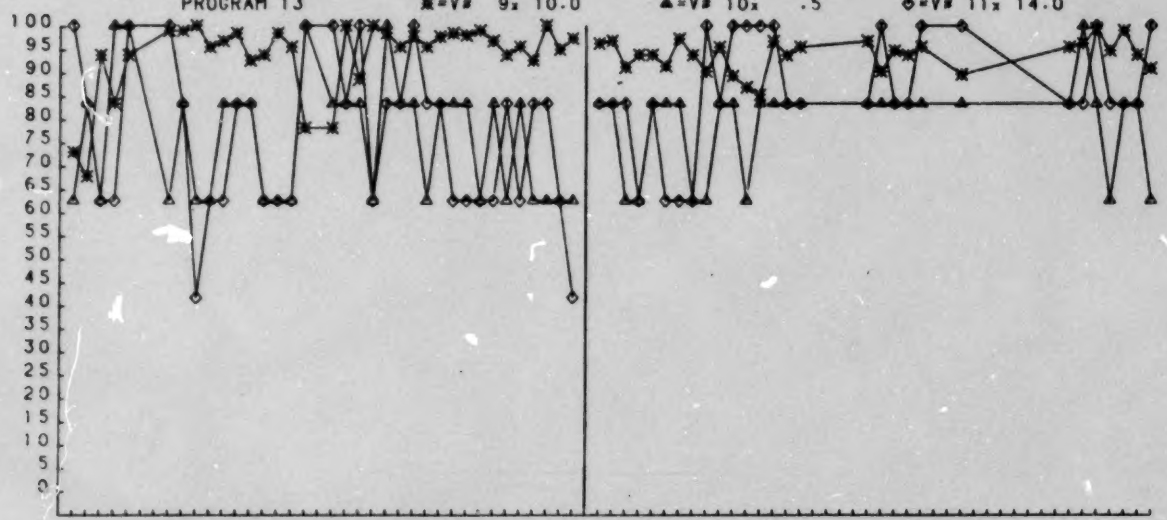

PROGRAM 12
✱=V# 9x 10.0
△=V# 10x .1
◇=V# 11x 14.0

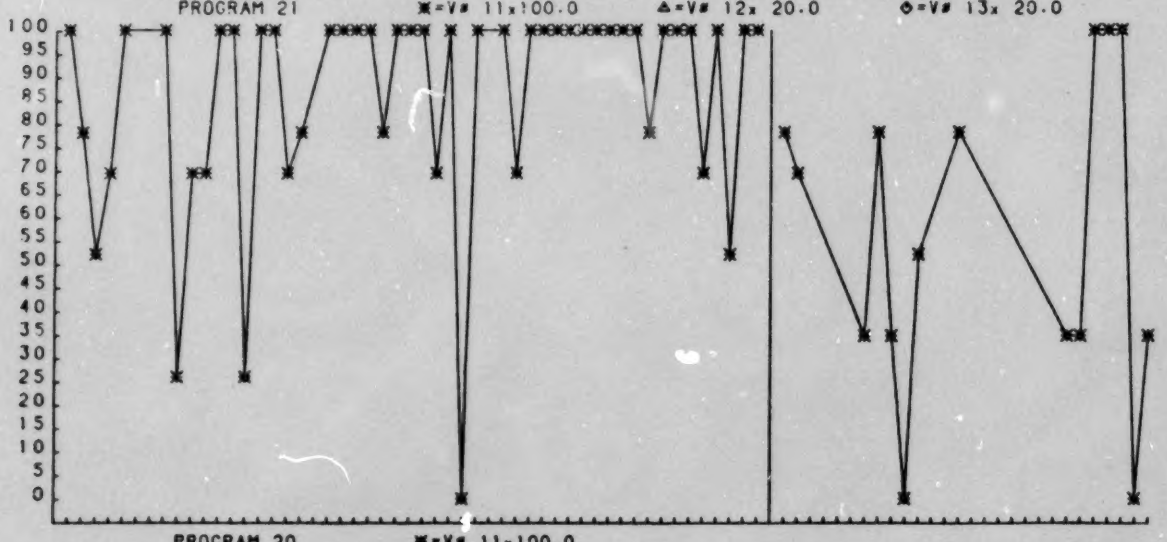

PROGRAM 13
✱=V# 9x 6.7
△=V# 10x .1
◇=V# 11x 14.0


PROGRAM 21
✱=V# 11x100.0
△=V# 12x 14.3
◇=V# 13x 14.3

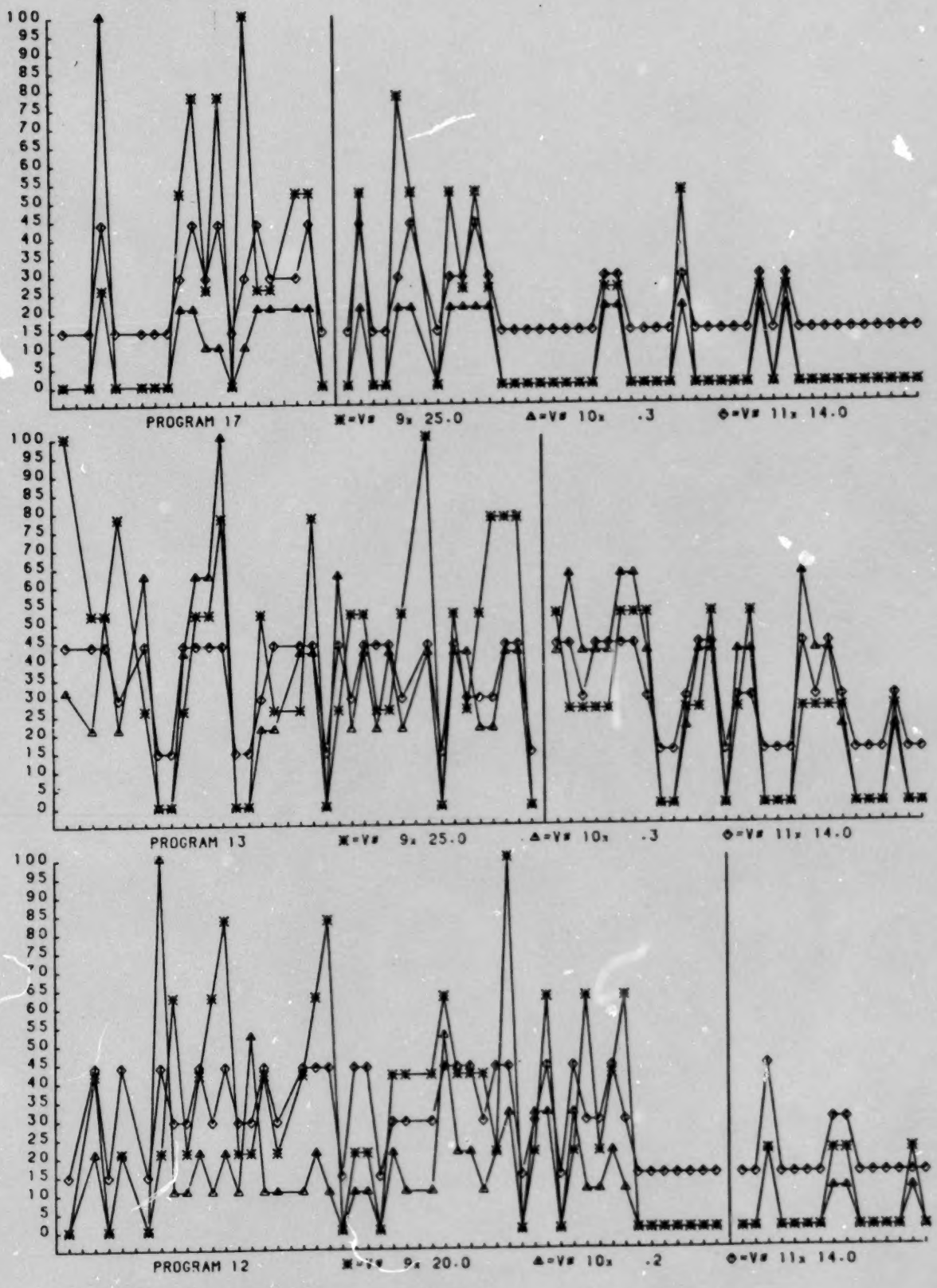
FAMILY 136

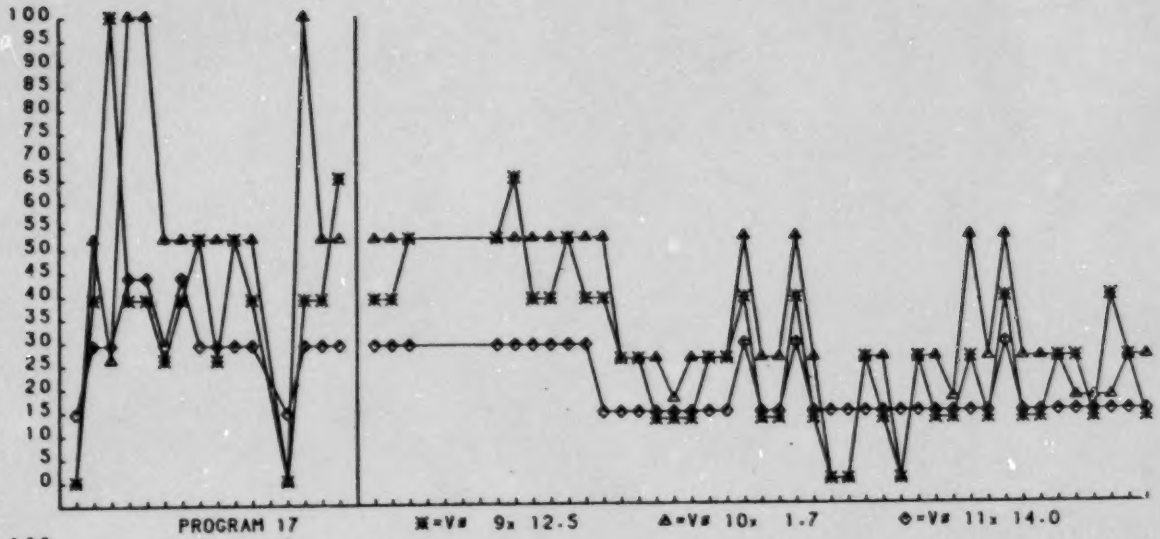

PROGRAM 13
*=V# 9x 10.0
▲=V# 10x .5
◇=V# 11x 14.0

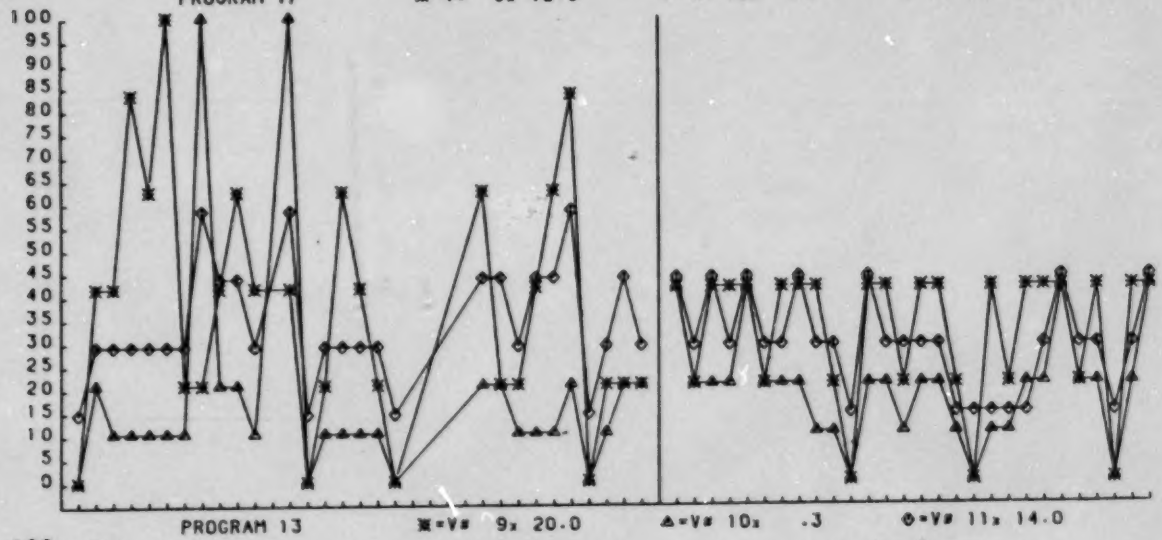

PROGRAM 21
*=V# 11x100.0
▲=V# 12x 20.0
◇=V# 13x 20.0

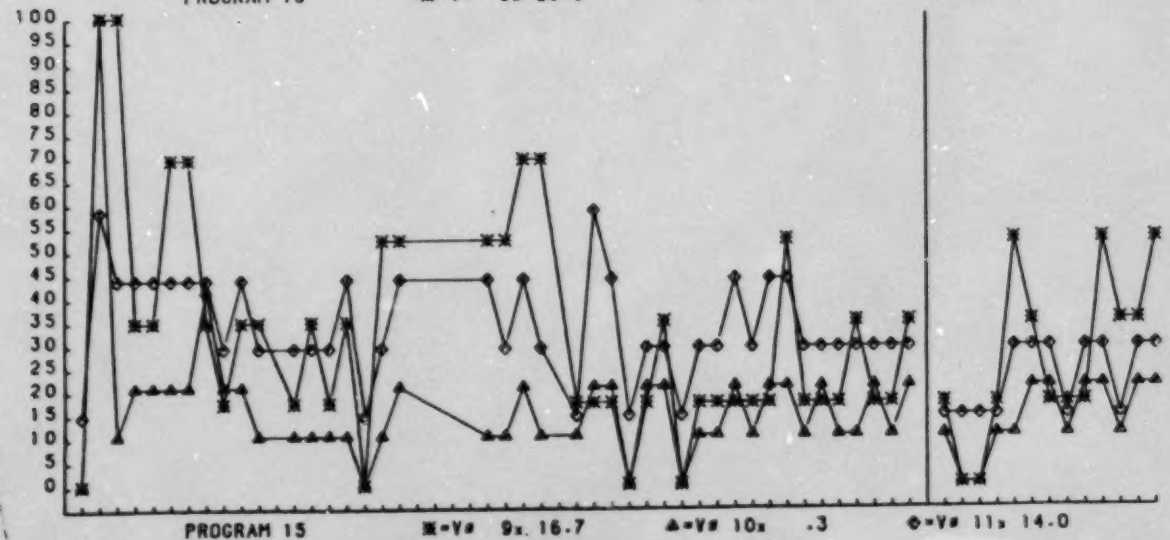

PROGRAM 20
*=V# 11x100.0

FAMILY 151

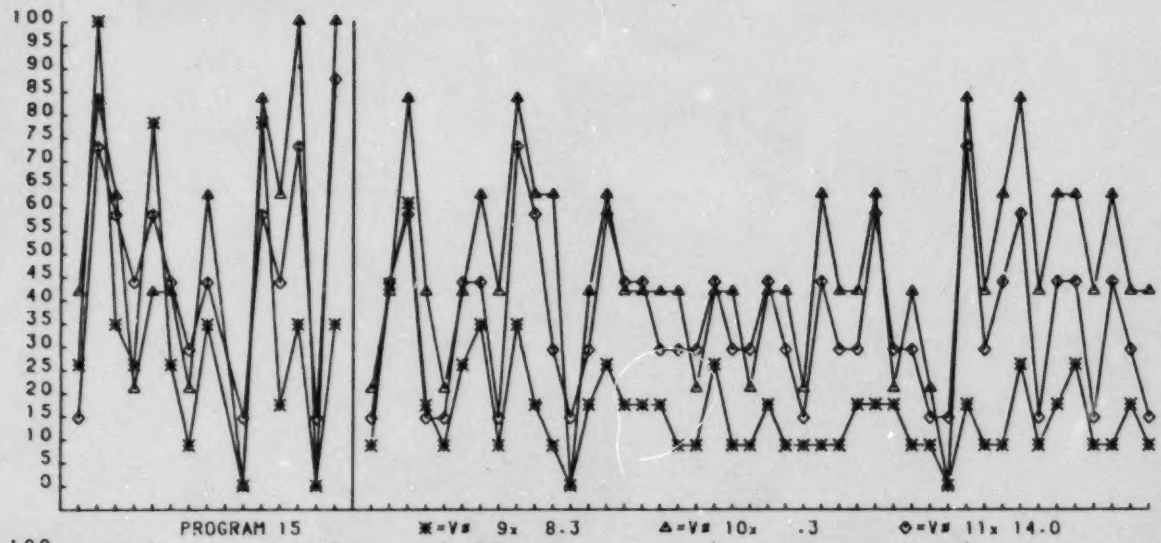

PROGRAM 17
PROGRAM 13
PROGRAM 12
FAMILY 152

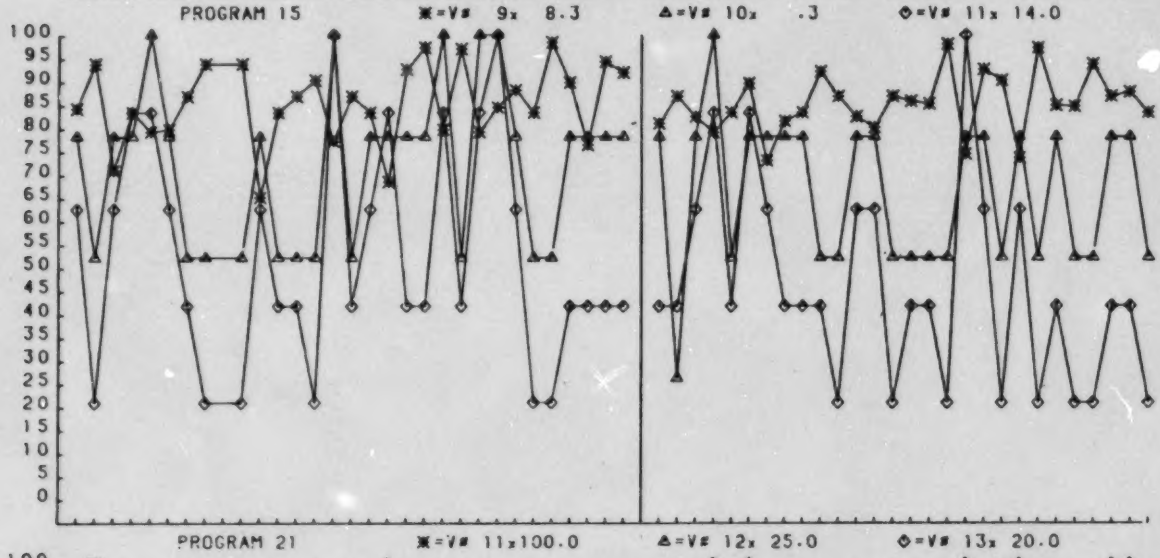

100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0
PROGRAM 17
*=V# 9: 12.5
▲=V# 10: 1.7
◆=V# 11: 14.0

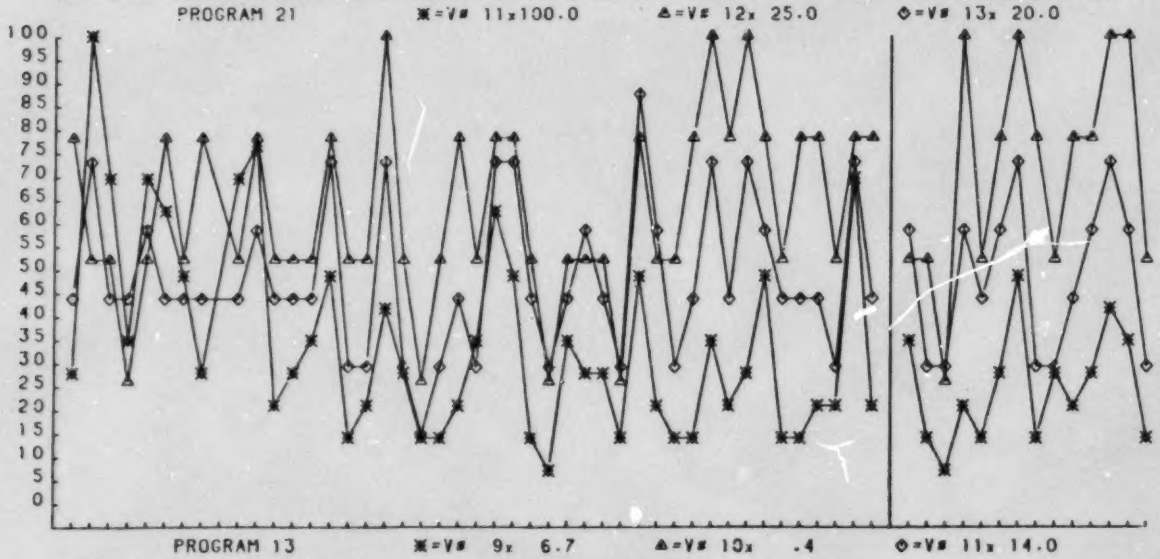

100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0
PROGRAM 13
*=V# 9: 20.0
▲=V# 10: .3
◆=V# 11: 14.0


100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0
PROGRAM 15
*=V# 9: 16.7
▲=V# 10: .3
◆=V# 11: 14.0

FAMILY 155


PROGRAM 15
✱=V# 9x 8.3
△=V# 10x .3
◇=V# 11x 14.0


PROGRAM 21
✱=V# 11x100.0
△=V# 12x 25.0
◇=V# 13x 20.0


PROGRAM 13
✱=V# 9x 6.7
△=V# 10x .4
◇=V# 11x 14.0

FAMILY 167

Consumer Evaluations

Six of the seven families who participated in treatment completed and mailed in a consumer evaluation. As each family used three programs, this gave 18 possible evaluations of specific programs. Eleven of these evaluations rated the overall effects as "good", seven as "neutral" and none as "bad" when these were the only three choices. Five of six families (parents) reported positive changes in themself (parent reporting) and one did not; four reported positive changes in their child while two did not; four reported positive changes in other family members and two did not; five reported positive changes in general and one did not. Five reported no "bad" or "negative" changes in themself, the child, or in general; one reported such negative changes. None reported negative changes in others. Four of the six indicated that they would use programs such as this again. Ratings of the usefulness, clarity, ease of use and pleasantness of use to the parent and the child generally suggested lower ratings for programs implemented last and higher ratings for programs implemented first. Ratings allegedly obtained from a limited number of chilren by their parents suggested that chilren were less positive about the programs than their parents; however, few children were said to have responded to the consumer evaluation.

CHAPTER SEVEN

Auxiliary Study

In a sub-study by Mullison (1982) a program for compliance related to but not similar to the following directions program was evaluated using an across subjects multiple baseline design. The written program used was a modified version of an early draft of the following directions program used in the project, combined with ideas of the author from the literature. A somewhat different observer recording system was also used. The subjects were three families with children aged 5,3 and 5 referred by a school psychologist due to non-compliant behavior. The third family dropped before starting treatment. In this family baseline rates of compliance were in what is usually considered the "normal" range.

With the two remaining families inter-observer reliabilities were obtained in 40% of all sessions, including baseline and treatment sessions. Percent agreement reliabilities for number of commands given and compliance ranged from 89% to 94% for both subjects. Pearson r correlation coefficients across all sessions (with reliability) averaged .98 for compliance and .9[illegible] for number of commands.

Subject 1 increased percent compliance from 50% during baseline to an average for all of treatment of 69%. In the final four treatment sessions percent compliance always was above 70%. The maximum possible improvement in compliance was 50%; 38% of the possible improvement was obtained.

Subject 2 increased percent compliance from 51% during baseline to an average for all of treatment of 70%. In the final four treatment

sessions percent compliance averaged just about 70%. The maximum possible improvement in compliance was 61%; 39% of the possible compliance was obtained.

For both subjects graphic analysis of data indicated that improvement coincided with the onset of the treatment program. However, with only two subjects the internal validity of this sub-study is weak.

APPENDIX

Funding Sources

Major funding was from the U.S. Office of Education, Grant No. G007910344.

Funding for computer data analysis and computer graphics was obtained from the Biomedical Research Committee of the University of Utah.

Additional funding was obtained from the Bureau of Educational Research, the Educational Psychology Department, and the Office of the Dean of the Graduate College of Education of the University of Utah.

Much of the indirect cost was contributed by the University of Utah.

Theses and Dissertations

McLoughlin, C.S. Evaluation of self-instructional parent-training byparent and others' reports. Unpublished doctoral dissertation, University of Utah, 1981.

Mullison, D.D. Self-instructional behavior therapy: the effectiveness of a parent training manual for reducing noncompliance. Unpublished master's thesis, University of Utah, 1982.

Convention Presentations, Including Those Accepted But Not Yet Presented

Endo, G.T. and Sloane, H.N. Untitled presentation on self-report measures in symposium on problems in parent training research, to be presented at the Association for Behavior Analysis, May, 1983, Milwaukee.

Sloane, H.N. and Endo, G.T. Results of self-instructional parent training. To be presented at the Association for Behavioral Analysis, May, 1983, Milwaukee.

Sloane, H.N. Invited presentation on parent training. Association for Behavior Analysis, May, 1981, Milwaukee.

McLoughlin, C.S., Sloane, H.N. and Endo, G.T. Research of parent training. Association for Behavior Analysis, May, 1981, Milwaukee.

Sloane, H.N. and Endo, G.T. The Utah self-instructional program. The Association for Behavior Analysis, May, 1980, Dearborn.

Sloane, H.N., Endo, G.T. and Miller, P. Real-time computer monitoring of single-subject research. The Association for Behavior Analysis, May, 1980, Dearborn.

Tentative, Projected Papers to be Prepared for Submission

1. Paper on self-instructional treatment of noncompliance by parents.

2. Paper on self-instructional treatment of fighting by parents.
3. Paper on self-instructional treatment of tantrums by parents.
4. Paper on self-instructional treatment of interrupting by parents.
5. Paper on self-instructional treatment of whining by parents.
6. Paper on self-instructional treatment of bedtime problems and room neatness
7. Paper on computer graphic program.
8. Paper on self-report measures.
9. Possible papers on relationships between or among parent accuracy in carrying out programs, outcomes, and reliability.